Marcin Anholcer

290

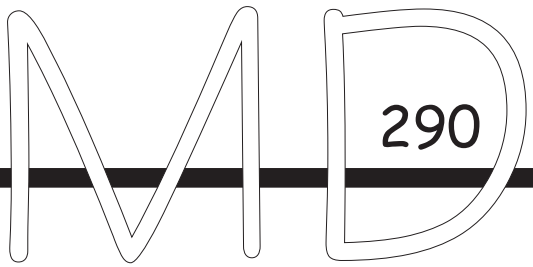
MATHEMATICS IN ECONOMICS AND MANAGEMENT

EXAMPLES AND EXERCISES

MATERIAŁY DYDAKTYCZNE

Wydawnictwo Uniwersytetu Ekonomicznego
w Poznaniu

UNIWERSYTET EKONOMICZNY
W POZNANIU



Marcin Anholcer

MATHEMATICS IN ECONOMICS AND MANAGEMENT

EXAMPLES AND EXERCISES

MATERIAŁY DYDAKTYCZNE

Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu

POZNAŃ 2015

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Editorial Board

Elżbieta Gołembska, *Danuta Krzemińska*, *Emil Panek*, *Wiesława Przybylska-Kapuścińska*, *Jerzy Schroeder* (secretary), *Ryszard Zieliński*, *Maciej Żukowski* (chairman)

Cover design
Weronika Rybicka

Production controller
DRUK-BAD Usługi Wydawnicze

© Copyright by Poznań University of Economics Poznań
2015

ISSN 1689-7412
ISBN 978-83-7417-836-5
eISBN 978-83-8211-030-2

POZNAŃ UNIVERSITY OF ECONOMICS PRESS
ul. Powstańców Wielkopolskich 16, 61-895 Poznań, Poland
phone +48 61 854 31 54, 61 854 31 55, fax +48 61 854 31 59
www.wydawnictwo-ue.pl, e-mail: wydawnictwo@ue.poznan.pl
postal adress: al. Niepodległości 10, 61-875 Poznań, Poland

Printed and bound in Poland by:
Poznań University of Economics Print Shop
ul. Towarowa 53, 61-896 Poznań, Poland
phone +48 61 854 38 06, +48 61 854 38 03

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

TABLE OF CONTENTS

Introduction 5

Chapter 1. Logic 7

1.1. Sentences 7
1.2. Quantifiers 15

Chapter 2. Sets 22

2.1. Basic Concepts 22
2.2. Cartesian Product 33
2.3. The Inclusion–Exclusion Principle 36

Chapter 3. Relations 40

3.1. Describing the Relations 40
3.2. Properties of the Relations 52
3.3. Orders and Preferences 58
3.4. Pareto Efficiency 63

Chapter 4. Matrices and Vectors 68

4.1. Basic Concepts. Matrix Algebra 68
4.2. Elementary Operations 73
4.3. Linear Independence of Vectors and Matrix Rank 81
4.4. Determinants 84
4.5. Matrix Inverse 88

Chapter 5. Linear Systems 94

5.1. Systems of Linear Equations 94
5.2. Systems of Linear Inequalities 101
5.3. Foundations of Linear Programming 110

Chapter 6. Basic Functions and Their Properties 119

6.1. Basic Concepts 119
6.2. Polynomials and Rational Functions 120
6.3. Exponential Function 127
6.4. Logarithmic Function 131

Chapter 7. Sequences and Series 137

7.1. Sequences and Their Limits 137
7.2. Arithmetic and Geometric Progressions 149
7.3. Sums and Series 151

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Chapter 8. Functions of One Variable 156

8.1. Limits 156
8.2. Derivatives 164

Chapter 9. Functions of Many Variables 175

9.1. Functions of Many Variables and Their Derivatives 175
9.2. Local Extrema 183
9.3. Conditional Extrema 190

Chapter 10. Integral Calculus 199

10.1. Indefinite Integrals 199
10.2. Definite and Improper Integrals 206

Chapter 11. Differential and Difference Equations 215

11.1. Differential Equations 215
11.2. Difference Equations 223

Chapter 12. Applications of Calculus in Economics 229

12.1. Differential Calculus 229
12.2. Important Functions 239

Chapter 13. Financial Mathematics 248

13.1. Compound Interest, Streams of Money and IRR 248
13.2. Loan Repayment Schedule 255

Chapter 14. Probability 261

14.1. Basic Concepts 261
14.2. Random Variable 267
14.3. Appendix: Cumulative Distribution of $N(0, 1)$ 282

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

INTRODUCTION

The book that you pick up has been developed for students of first year courses in International Relations, Finance and Management. It has been designed to help you – even those who learned mathematics in high school only at the elementary level – to pass the exams in mathematics (or any other similar courses).

The book is prepared in a way that allows you to learn by yourself. Hence, great emphasis was placed on the examples that dominate in this book. They are resolved "step by step" to make it easy to understand them.

I also briefly discussed some of the issues that should be discussed in high school, but often cause some trouble: solving polynomial and rational equations and inequalities, basic properties of exponential and logarithmic functions.

The book consists of fourteen chapters, which can be divided into four conventional parts. The **first part**, consisting of Chapters 1, 2 and 3, introduce you to the basic issues related to **logic and set theory**. It has been discussed, inter alia, issues such as the sentence algebra, quantifiers, set theory, and relations. In the **second part** (Chapters 4 and 5) I discussed the basics of **linear algebra**. You will learn from it how to apply arrays, vectors, determinants and linear systems in economics. In the largest **third part** (Chapters 6–13), I discussed the basic concepts of **calculus** and their **economic applications**. So here you will find the methods of determining the limits and derivatives, finding the extrema, and solving the differential and difference equations. Applications have been collected in Chapters 12 and 13 – here you will find, among others, the information about marginal analysis and foundations of financial mathematics. The last, **fourth part** consists of one Chapter (14). I discussed there the foundations of **probability theory**, including most important from the point of view of your future work, random variables.

Each chapter consists of at least two sections, almost each section consists of four parts. First of them is a short theoretical introduction (it is enough for understanding the examples and solving the exercises, but it is not exhausting the topic). Almost each theoretical part ends with a short description of how to use **WolframAlpha®** to solve the examples of respective kind. WolframAlpha® is a free web application derived by Wolfram (http://www.wolfram.com/). You can find the application on the page http://www.wolframalpha.com/. The usage of WolframAlpha® is very simple (you only type the commands and click the "=" button on the right). You can also access easily the huge number of self-explanatory examples by clicking "Examples". Finally, if you want to use more advanced features of the application (for example the "step by step" solutions), there is a possibility to buy a premium license.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The book was created on the base of my professional experience. It is a extended version of the Polish book *Matematyka w ekonomii i zarządzaniu w przykładach i zadaniach* (MD 360, Poznań University of Economics, 2020).
I hope that the book that you hold in your hand, will expand your mathematical knowledge and help you in passing the math exam.

Marcin Anholcer

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 1
LOGIC

1.1. SENTENCES

THEORY IN A NUTSHELL

Logical statement (or sentence) is any declarative sentence, whose truthfulness may be verified. One can distinguish simple statements, which cannot be reduced and the compound (or complex) statements, consisting of the simple ones.
Simple statements are denoted by lowercase letters: p, q, r, … The logical value (truth value, Boolean value) is denoted by the Greek letter ν (nu). It takes always one of two values: 0, when the statement is false, and 1, when it is true. We write it in the following way:

$$\nu(p) = \begin{cases} 1, & \text{if } p \text{ is true,} \\ 0, & \text{otherwise.} \end{cases}$$

As mentioned before, it is possible to construct compound statements based on the simple ones. What we need are the operations. Five of them are described below.
Negation is a one-argument (unary) operation. We denote it by "$\neg$"[1] and read simply "not". The so-called **truth table** for negation (i.e. the list of all the possible values of it depending on the truth values of the initial simple statement) is presented in Table 1.1. As you can see, the value of $\neg p$ is always different that the value of p.

Table 1.1. Truth table of the negation

p	$\neg p$
0	1
1	0

Four remaining operations are conjunction ("$p \wedge q$", we read it "p and q"), disjunction ("$p \vee q$", we read it "p or q"), implication ("$p \Rightarrow q$", we read it "if p, then q") and equivalence ("$p \Leftrightarrow q$", we read it "p if and only if q"). The **conjunction** of two statements is true only if both of these statements are true. The **disjunction** is true, if at least one of them is true (note that it means that

[1] You can also meet „~".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

also they both can be true). The **implication** is true, when the first statement (predecessor) is false or the second statement (consequent) is true. So it is false only if the predecessor is true and the consequent is false (you can explain yourself in a way that with real benchmarks you cannot draw false conclusions). The **equivalence** is true if both statements are equivalent, that is, when they have the same truth value. The truth values of conjunction, disjunction, implication and equivalence are listed in Table 1.2.

Table 1.2. Truth values of conjunction, disjunction, implication and equivalence

p	q	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

Similarly as in the case of operations on numbers, also in the case of logical operations there is the predefined order of performing the operations. The first one is the negation, then the conjunction and disjunction, finally the implication and equivalence. For instance in the statement "$[(\neg p) \vee q] \Leftrightarrow (p \wedge q)$" we can skip all the brackets and it won't change neither its meaning nor the logical value. In turn the notation "$p \wedge q \vee r$" is ambiguous, since conjunction and disjunction have same priority. Thus the meaning of "$p \wedge (q \vee r)$" is not the same as the meaning of "$(p \wedge q) \vee r$".

Note. In order to simplify the notation, we usually do not use the function v to denote the value of the statement. Instead of it we simply write "$p \Leftrightarrow 0$" when p is false and "$p \Leftrightarrow 1$" when it is true. Similarly, we will skip the quotation marks in the remainder of this book.

Definition 1.1. Tautology
Tautology is a logical sentence that is always true, regardless from the values of simple sentences that create it.

To check if the sentence is logical tautology, it is necessary to check it's value for every possible combination of the values of simple sentences that create it. In order to do that, we use the so-called truth tables, discussed in more detailed way in the section "EXAMPLES".
In order to implement the truth table in **WolframApha®**, we use the command "truth table". The sample symbols for the operations are: "~" (negation), "||" (disjunction), "&&" (conjunction), "=>" (implication) and "<=>" (equivalence).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

For instance, in order to check the sentence $(\neg p \vee q) \Leftrightarrow (p \wedge q)$, we will use "truth table (~p||q)<=>(p&&q)".

Table 1.3 contains the list of the most famous and frequently used tautologies.

Table 1.3. Important tautologies

Tautology	Name
$\neg\neg p \Leftrightarrow p$	Double negation law
$p \Rightarrow p$	Implication law
$p \Leftrightarrow p$	Equivalence law
$p \wedge \neg p \Leftrightarrow 0$	Contradiction (1st negation law)
$p \vee \neg p \Leftrightarrow 1$	Excluded middle (2nd negation law)
$p \vee q \Leftrightarrow q \vee p$	Commutative laws
$p \wedge q \Leftrightarrow q \wedge p$	
$p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$	Associative laws
$p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$	
$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$	Distributive laws
$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$	
$p \wedge 1 \Leftrightarrow p$	Absorption laws
$p \vee 1 \Leftrightarrow 1$	
$p \wedge 0 \Leftrightarrow 0$	
$p \vee 0 \Leftrightarrow p$	
$\neg (p \vee q) \Leftrightarrow (\neg q \wedge \neg p)$	De Morgan's laws
$\neg (p \wedge q) \Leftrightarrow (\neg q \vee \neg p)$	

EXAMPLES

Example 1
Prove, that the double negation law is a tautology.

Solution
In the double negation law there is only one simple sentence, and so we have to consider only two cases, $p \Leftrightarrow 0$ and $p \Leftrightarrow 1$.
For each of them we determine the value of the expression $\neg\neg p$. We do this in two stages. First, we determine the value of the expression $\neg p$. We use for this purpose the information contained in the Table 1.1. The results are presented in the Table 1.4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Table 1.4. Solving the example – step 1

p	$\neg p$
0	1
1	0

In the second step, we determine the value of the expression $\neg\neg p$. Also in this case we use the information from the Table 1.1, bearing in mind, however, that this time the input data are the values of $\neg p$. In the first row, the expression $\neg p$ takes the value 0, so the value of $\neg\neg p$ is 1. In the second line the other way around. The results are listed in the Table 1.5.

Table 1.5. Solution of the example

p	$\neg p$	$\neg\neg p$
0	1	0
1	0	1

It is time to determine the values of the whole sentence. This has the form of equivalence, the left side is $\neg\neg p$ (third column of Table 1.5), the right side – p (first column of Table 1.5). As it follows from the information provided (for example) in Table 1.2, the equivalence is true if both component sentences have the same truth value. In the first case both $\neg\neg p$ and p are false, and therefore have the same truth value. Thus, the logical value of $\neg\neg p \Leftrightarrow p$ is 1. Similarly, in the second case, the two formulae ($\neg\neg p$ and p) are true. The final results are presented in Table 1.6.

Table 1.6. Solution of the example

p	$\neg p$	$\neg\neg p$	$\neg\neg p \Leftrightarrow p$
0	1	0	1
1	0	1	1

As you can see, in both cases (i.e. in all the possible cases) the logical value of the formula is 1, so it is a tautology.

Example 2
Check if the formula $(p \vee q) \wedge (p \Leftrightarrow r) \Rightarrow (q \wedge r)$ is a tautology.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

Since there are three simple statements in the formula (p, q, r), we have to consider three cases: two for p, each of them corresponding to two cases for q (giving a total of four possibilities), and each of the resulting possibilities corresponds to the two possible values for r. Generally, if a formula consists of n simple statements, then you need to consider 2^n different combinations of Boolean values. In order to list all of them, you should print them in a systematic way. You can use, for example, the following method. First, we set the value of a sentence in such a way that we alternately enter 0 and 1 then setting the value of the second sentence, we enter two same numbers: two times 0 two times 1, and so on. For every following sentence we group 0s and 1s in the blocks which are twice as long as the previous ones (i.e., third sentence – four same values in each block, fourth – eight and so on.). The effect of such a procedure for the current of example are presented in Table 1.7. The values of r were entered alternately, the values of q in pairs, and the values of p – in quadruples.

Table 1.7. Solution of the example

p	q	r
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

Now we can proceed to the main part of the solution. First, for each of the discharged cases, we determine the logical values of expressions $p \vee q$, $p \Leftrightarrow r$ and $q \wedge r$. Let us begin with $p \vee q$. In the first row, the Boolean values of the statements p and q are 0. From the information in the Table 1.2 it follows that in such a case, the value of the disjunction is also 0. We have same situation in the second row. In the rows 3 and 4, p has value 0, and q has value 1. Looking again into the Table 1.2, we find that in such a situation (one statement false, the second true), the alternative is true. Doing so for all rows, we finally obtain the logical values of $p \vee q$ listed in the Table 1.8.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Table 1.8. Solution of the example

p	q	r	$p \vee q$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

We proceed similarly with the formulae $p \Leftrightarrow r$ and $q \wedge r$. Just remember to read the values of their elements from appropriate columns. Let us consider for instance the row 4 (p is false, q and r are true). In such a case the value of $p \Leftrightarrow r$ is 0 (the value of p is 0, the value of r is 1, so they are different), and the value of $q \wedge r$ is 1 (both q and r are true, so it follows from the Table 1.2 that their conjunction is also true). The results are listed in the Table 1.9.

Table 1.9. Solution of the example

p	q	r	$p \vee q$	$p \Leftrightarrow r$	$q \wedge r$
0	0	0	0	1	0
0	0	1	0	0	0
0	1	0	1	1	0
0	1	1	1	0	1
1	0	0	1	0	0
1	0	1	1	1	0
1	1	0	1	0	0
1	1	1	1	1	1

The next step is evaluating the left side of the investigated formula, i.e. $(p \vee q) \wedge (p \Leftrightarrow r)$. Again we use the contents of the Table 1.2, bearing in mind that the calculations use the logical values of the expressions $(p \vee q)$ and $(p \Leftrightarrow r)$. For example, in the first row $(p \vee q)$ has value 0, while $(p \Leftrightarrow r)$ has value 1, so the value of their conjunction will be 0. The results are presented in the Table 1.10.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Table 1.10. Solution of the example

p	q	r	$p \vee q$	$p \Leftrightarrow r$	$q \wedge r$	$(p \vee q) \wedge (p \Leftrightarrow r)$
0	0	0	0	1	0	0
0	0	1	0	0	0	0
0	1	0	1	1	0	1
0	1	1	1	0	1	0
1	0	0	1	0	0	0
1	0	1	1	1	0	1
1	1	0	1	0	0	0
1	1	1	1	1	1	1

In the last step, we determine the logical values of the whole formula, bearing in mind that it has a form of implication, with the components $(p \vee q) \wedge (p \Leftrightarrow r)$ and $q \wedge r$, and so the logical values of those expressions must be used in the calculation. For example, in the third row, they are respectively 1 and 0 (note the order!), So the implication is false (Boolean value 0), and in the fourth row 0 and 1, so the implication is true (Boolean value 1). The final results are presented in the Table 1.11.

Table 1.11. Solution of the example

p	q	r	$p \vee q$	$p \Leftrightarrow r$	$q \wedge r$	$(p \vee q) \wedge (p \Leftrightarrow r)$	$(p \vee q) \wedge (p \Leftrightarrow r) \Rightarrow (q \wedge r)$
0	0	0	0	1	0	0	1
0	0	1	0	0	0	0	1
0	1	0	1	1	0	1	0
0	1	1	1	0	1	0	1
1	0	0	1	0	0	0	1
1	0	1	1	1	0	1	0
1	1	0	1	0	0	0	1
1	1	1	1	1	1	1	1

In conclusion, after examining the eight possible cases, we find that in two of them the formula is false. This means that it is not a tautology (recall – tautology is a sentence always that is true).

Example 3

It is known that the formulae $p \vee \neg q$ and $p \Rightarrow q$ are true. What can we say about the statements p and q?

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

So far we determined the logical values of the formulae, knowing the values of the simple statements being their elements. Now we have to do the opposite. But actually it all boils down to the same scheme. We start from writing all the possible combinations of simple statements (in our example: p and q), then on that basis we determine the values of the resulting formulae. Only the conclusions we draw in a different way. So let's determine the values of the formulae. The calculation results are contained in the Table 1.12 (note that due to the presence of two simple statements we have to consider $2^2 = 4$ cases).

Table 1.12. Solution of the example

p	q	$\neg q$	$p \vee \neg q$	$p \Rightarrow q$
0	0	1	1	1
0	1	0	0	1
1	0	1	1	0
1	1	0	1	1

Time to conclude. Since we know that $p \vee \neg q$ and $p \Rightarrow q$ are true, we have to find the lines corresponding to the respective values (1 in the last two columns) in Table 1.12. These are the rows 1 and 4. In the first two columns we read that corresponds to the values of statements p and q equal to 0 and 0 (first row) or 1 and 1 (fourth row).

So finally given formulae are true when both statements p and q are false or both are true.

EXERCISES

1. Check if the following formulae are tautologies:
   a) $p \Rightarrow p$;
   b) $p \Leftrightarrow p$;
   c) $p \wedge \neg p \Leftrightarrow 0$;
   d) $p \vee \neg p \Leftrightarrow 1$;
   e) $p \wedge 1 \Leftrightarrow p$;
   f) $p \vee 1 \Leftrightarrow 1$;
   g) $p \wedge 0 \Leftrightarrow 0$;
   h) $p \vee 0 \Leftrightarrow p$;
   i) $p \vee q \Leftrightarrow q \vee p$;
   j) $p \wedge q \Leftrightarrow q \wedge p$;
   k) $\neg (p \vee q) \Leftrightarrow (\neg q \wedge \neg p)$;
   l) $\neg (p \wedge q) \Leftrightarrow (\neg q \vee \neg p)$;
   m) $p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$;
   n) $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

o) $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$;
p) $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$;
q) $(p \vee q) \wedge (p \vee r) \Rightarrow p \wedge q \wedge r$;
r) $(p \vee q) \wedge (p \wedge r) \Rightarrow p \wedge q \wedge r$;
s) $(p \wedge q) \vee (r \wedge s) \Rightarrow (p \vee r) \wedge (q \vee s)$;
t) $(p \vee r) \wedge (q \vee s) \Rightarrow (p \wedge q) \vee (r \wedge s)$.

2. Find the logical values of p, q, r, if we know that the following formulae are true:
   a) $p \Rightarrow q, p \vee q$;
   b) $(p \vee q) \wedge r, r \Rightarrow \neg p$;
   c) $p \vee q \Rightarrow p \wedge q$;
   d) $p \vee q \vee r \Rightarrow p \wedge q \wedge r$;
   e) $p \Rightarrow q \vee r, q \Rightarrow r$.

SOLUTIONS

1. a) – p) yes; q), r) no; s) yes; t) no.
2. Logical values:
   a) p arbitrary, q true;
   b) q and r true, p false;
   c) both true or both false;
   d) all true or all false;
   e) all three formulae false or r true, and two remaining arbitrary.

1.2. QUANTIFIERS

THEORY IN A NUTSHELL

Predicate is an expression whose logical value depends on certain of its parts, called **variables**. An example of predicate is the expression "$x^2 = 49$". You cannot say that it is true or it is false, until you determine the value of the variable x. If we know that $x = -7$, then the above expression is true. But if $x = 4$, this expression is false. As you can see, only after establishing the value of the variable the Boolean value of the expression can be evaluated. There is absolutely no need to give a specific value of a variable, it can be described in a less precise manner, for example: "for every x being a real number $x^2 = 49$" (which is not true, just check the above mentioned value of $x = 4$) or "there exists an integer x such that $x^2 = 49$" (which in turn is true, as an example check $x = -7$).
Expressions having such a form ("for every x the property $P(x)$ occurs", "there is an x such that the property $P(x)$ occurs") are very helpful in defining the mathematical concepts (and not only). They are stored using special symbols called quantifiers. There are several quantifiers, but we are going to use only two

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

of them: **universal quantifier** corresponding to the phrase "**for all**" and **existential quantifier** corresponding to the expression "**exists**". They are given in the Table 1.13:

Table 1.13. Quantifiers

Type	How we read it?	Symbol*
Universal	"For each …"	$\forall$
Existential	"Exists …"	$\exists$

*Sometimes you can also meet the following notation: "$\bigwedge$" for the universal quantifier and "$\bigvee$" for the existential quantifier.

Symbols shown in Table 1.13 are used in such a way that the range of a variable is given below the symbol, and the predicate is written on the right, or both the scope of the variable, as well as the predicate are written on the right, separated by a colon. For example the sentence "For each real number x, $x^2 \geq 0$" we can write at least in two ways – this way:

$$\forall x \in R : x^2 \geq 0,$$

or this way:

$$\bigvee_{x \in R} x^2 \geq 0.$$

Sentences written by logic quantifiers are subject to certain dependencies. In particular, if $P(x)$ denotes any predicate with variable x, then the two following properties, called De Morgan's laws for quantifiers, are true:

$$\left(\bigvee_x \neg P(x) \right) \Leftrightarrow \neg \left(\bigexists_x P(x) \right),$$

$$\left(\bigexists_x \neg P(x) \right) \Leftrightarrow \neg \left(\bigvee_x P(x) \right).$$

As you can see, De Morgan's laws say that after moving the negation inside the predicate, we have to change its type (if you see the analogy to the De Morgan's laws logical statements?).
Quantifiers of the same type can be stored in any order. It does not matter in what order we write some universal quantifiers or the existential ones. It is in turn important to keep the order while writing the quantifiers of different types. Whether it appears first existential quantifier, then an universal one, or vice

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

versa has an impact on the meaning and value of the statement. This can be easily seen by examining the first example.

EXAMPLES

Example 1

Read the following expressions and find their Boolean values:

$$\forall_{x \in R} \exists_{y \in R} x + y = 0,$$

$$\exists_{y \in R} \forall_{x \in R} x + y = 0.$$

Solution

At the beginning of the first expression there is the universal quantifier with variable x, which should read: "for every number x belonging to the set of real numbers". Another element of the expression is the existential quantifier with variable y, so read it: "there is a number y that belongs to the set of real numbers such that". Finally, the sentence sounds like "for every number x belonging to the set of real numbers, there is a number y that belongs to the set of real numbers such that x plus y is equal to 0".

Is this true? Yes. Why? Because for each x it is enough to take $y = -x$. Then indeed $x + y = x + (-x) = 0$.

And what would happen if we switched the quantifiers? The sentence would read as follows: "there exists a real number y such that for each real number x the sum of x and y is 0". This is not true – there is no such number, which would give 0 after combining with **any** real number.

As you can see, the order of the quantifiers really matters.

At the end, pay attention to the fact that once we have used the term "real number x", and once "x belonging to the set of real numbers". Both are correct. You can alternatively use any other, similarly sounding expressions, it is only important to preserve the sense (x is the number from a specific set).

Example 2

Write the following sentence using quantifiers: "for any two natural numbers m and n, their product is a nonnegative number".

Solution

The expression „for any" has in this context exactly the same meaning as „for each". So we have to use two universal quantifiers – one for m one for n, or one universal quantifier for m and n. In the first case the formula looks like this:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\forall_{m\in N}\forall_{n\in N} mn \geq 0.$$

In the second case we can use any of the following notations:

$$\forall_{m\in N, n\in N} mn \geq 0,$$

$$\forall_{\substack{m\in N\\ n\in N}} mn \geq 0.$$

As both numbers are from the same set, we can also use the simplified notation:

$$\forall_{m,n\in N} mn \geq 0.$$

Example 3
Check whether the given statement is true. If not, save its negation using De Morgan's laws.

$$\exists_{x\in N}\forall_{y\in R} x + y > 0.$$

Solution
This sentence says that there is a natural number x such that after summing up with any real number y we get always a positive number. So it is not true, because regardless of the choice of x, it is sufficient to take $y = -x$ to obtain $x + y = x + (-x) = 0$. Let us construct the negation then:

$$\neg\exists_{x\in N}\forall_{y\in R} x + y > 0.$$

Let us apply De Morgan's law to the first quantifier. As a result of its use, the existential quantifier turns into the general one and the symbol of negation moves to the right. We obtain the following form of the expression:

$$\forall_{x\in N}\neg\forall_{y\in R} x + y > 0.$$

Now we apply De Morgan's law to the second quantifier. Its use makes the universal quantifier turns into the existential one, and the symbol of negation moves further to the right. We get:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\bigvee_{x\in N}\bigexists_{y\in R}\neg\, x+y>0.$$

Finally, let us write the expression under the quantifier in more accessible way. To do this, we change the inequality. Remember that if you want to express something opposite to what expresses the inequality "$>$" is not enough to turn it! You also have to take into account the situation of equality of both sides, and so the use of "$\leq$" is necessary (similarly the change of the Boolean value of a sentence requires the conversion of inequality "$<$" to "$\geq$"). So we get finally the following form:

$$\bigvee_{x\in N}\bigexists_{y\in R} x+y\leq 0.$$

EXERCISES

1. Read the following expressions and find their logical values:
   a) $\bigvee_{x\in R} x^2>0;$
   b) $\bigvee_{x\in R} x^2\geq 0;$
   c) $\bigexists_{x\in R} x^2>0;$
   d) $\bigexists_{x\in R} x^2<0;$
   e) $\bigvee_{x,y\in R} x^2+y^2+1>0;$
   f) $\bigexists_{x,y\in R} x+y<-5;$
   g) $\bigvee_{m\in N}\bigexists_{x\in R} m^x>9932;$
   h) $\bigexists_{m,n\in N}\bigvee_{x\in R} mn=mx;$
   i) $\bigvee_{x\in R}\bigexists_{m\in N}(x-m)^2<0;$
   j) $\bigvee_{\varepsilon>0}\bigexists_{\delta>0}\bigvee_{x\in R}(|x-x_0|<\delta\Rightarrow|f(x)-g|<\varepsilon).$
2. Write the following sentences using the symbols and find their logical values:
   a) Each natural number is greater than 0.
   b) Every real number is negative.
   c) There are some negative natural numbers.
   d) There are no real numbers with negative squares.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

e) There is a natural number such that there is no real number that would be greater than it.
f) Each time when a real number is negative, then also its square is negative.
g) If there is a natural number greater than 0, then there exists a real number lower than 0.
h) Only the squares of natural numbers are natural numbers.
i) Squares of some real numbers, not being natural numbers, are natural numbers.
j) For each number ε greater than 0 there exists a number M such that for each natural number m the following is true: if m is greater than M, then the absolute value of the difference between the value of function f at point m and the number g is lower than ε.

3. Write the negations of the expressions from exercise 1, using the De Morgan's formulae.

SOLUTIONS

1. For example:
   a) For each real number x its square is greater than 0. False.
   b) For each real number x its square is greater than or equal to 0. True.
   c) There is a real number x such that its square is greater than 0. True.
   d) There exists a real number x such that its square is lower than 0. False.
   e) For each pair of real numbers x and y the expression $x^2 + y^2 + 1$ is positive. True.
   f) There is a pair of real numbers x and y such that their sum is lower than negative 5. True.
   g) For each natural number m there is a real number x such that m to the power x is greater than 9932. False[2].
   h) There is a pair of natural numbers m i n such that for each real number x the product of m and n equals to the product of m and x. True.
   i) For every real number x there is a natural number m such that the square of the difference between x and m is negative. False.
   j) For each ε greater than 0 there is a δ greater than 0 such that for every real number x the following is true: if the absolute value

[2] In this book we assume that 0 is a natural number. The expression 0^x equals to 0, when $x > 0$ and is not defined for $x \leq 0$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

of the difference between x and x_0 is lower than δ, then the absolute value of the difference between the value of the function f at x and the number g is lower than ε. We cannot say whether it is true or not (we know nothing about x_0, g and f).

2. Sample notation:
   a) $\underset{m\in N}{\forall} m > 0$. False.
   b) $\underset{x\in R}{\forall} x < 0$. False.
   c) $\underset{m\in N}{\exists} m < 0$. False.
   d) $\neg \underset{x\in R}{\exists} x^2 < 0$. True.
   e) $\underset{m\in N}{\exists} \neg \underset{x\in R}{\exists} x > m$. False.
   f) $\underset{x\in R}{\forall}\left(x < 0 \Rightarrow x^2 < 0\right)$. False.
   g) $\left(\underset{m\in N}{\exists} m > 0\right) \Rightarrow \left(\underset{x\in R}{\exists} x < 0\right)$. True.
   h) $\underset{m}{\forall}\underset{n}{\forall}(n = m^2 \wedge n \in N \Rightarrow m \in N)$. False.
   i) $\underset{x\in R}{\exists}\underset{n\in N}{\exists}(\neg x \in N \wedge x^2 = n)$. True.
   j) $\underset{\varepsilon>0}{\forall}\underset{M\in N}{\exists}\underset{m\in N}{\forall}(m > M \Rightarrow |f(m) - g| < \varepsilon)$. We cannot say whether it is true or not (we know nothing about x_0, g and f).

3. The negations:
   a) $\underset{x\in R}{\exists} x^2 \le 0$;
   b) $\underset{x\in R}{\exists} x^2 < 0$;
   c) $\underset{x\in R}{\forall} x^2 \le 0$;
   d) $\underset{x\in R}{\forall} x^2 \ge 0$;
   e) $\underset{x,y\in R}{\exists} x^2 + y^2 + 1 \le 0$;
   f) $\underset{x,y\in R}{\forall} x + y \ge -5$;
   g) $\underset{m\in N}{\exists}\underset{x\in R}{\forall} m^x \le 9932$;
   h) $\underset{m,n\in N}{\forall}\underset{x\in R}{\exists} mn \ne mx$;
   i) $\underset{x\in R}{\exists}\underset{m\in N}{\forall}(x-m)^2 \ge 0$;
   j) $\underset{\varepsilon>0}{\exists}\underset{\delta>0}{\forall}\underset{x\in R}{\exists}(|x - x_0| < \delta \wedge |f(x) - g| \ge \varepsilon)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 2
SETS

2.1. BASIC CONCEPTS

THEORY IN A NUTSHELL

Set is most often classified as one of the **primitive notions**, that is undefined. For our purposes, we can characterize set as a collection of **unique and not ordered elements**[3].
We denote the sets by capital letters: A, B, C, …, and their components (called **members**) – by the lowercase letters. The fact that the element x belongs to the set A, is written as: $x \in A$, while the fact that it doesn't: $x \notin A$.
The number of members of A is denoted by $|A|$ and is called the cardinality of A. If it is finite, sometimes it is denoted by $n(A)$.
Sets may be described in various ways. In this book, we will be mostly interested in finite sets and sets of numbers. Therefore, to represent the members of a set we will use:

- enumeration of all the members of the set, for example $A = \{1, 2, 3, 4\}$;
- notation using formulae, inequalities, equations, etc., for example: $A = \{x : x \in N \wedge x \geq 1 \wedge x \leq 4\}$ (":" – colon – means "such that"; given expression mean thus "A is the set of x such that x is a natural number, x is not lower than 1 and x is not greater than 4"; note that the set defined this way is exactly the same as the set from the previous bullet point);
- notation using intervals (see later in this chapter), for example: $A = [1, 4]$ (observe that this time the set is not same as before, as for instance 3.5 is its member).

The **equality** of two sets A and B means that they have exactly the same members. We write it as $A = B$. More formally, equality of two sets can be defined as follows:

$$A = B \Leftrightarrow \left(\bigvee_x x \in A \Leftrightarrow x \in B \right).$$

The **inclusion** of two sets A and B means that all the members of one set are also the members of another one. Inclusion of A in B is denoted by $A \subset B$ or $A \subseteq B$. In such a case we say that A **is subset of** B. If A is not subset of B, then we

[3] In the case of repetitions of elements we are talking about multisets. In the case of ordered elements we are talking about sequences. This last concept is discussed in detail later in the book.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

denote it by $A \not\subset B$. More formally, inclusion of A in B can be defined as follows:

$$A \subset B \Leftrightarrow \left(\bigvee_x x \in A \Rightarrow x \in B \right).$$

Note: inclusion of one set in another does not mean that they are equal! In order to obtain equality, mutual inclusion must occur. As an example, let us consider the following sets: $A = \{1, 2, 3\}$ and $B = \{1, 2\}$. In this case $B \subset A$, because all the members of B (that is "1" and "2") belong also to A. On the other hand "3" belongs to A and does not belong to B, so $A \not\subset B$.

Remember that if $A = B$, then both inclusion relations must occur: $A \subset B$ and $B \subset A$. In particular, for any set A it is true that $A \subset A$.

A special type of set is the **empty set**, denoted by the symbol $\varnothing$. Empty set **has no members**. It follows in particular that for any set A the condition $\varnothing \subset A$ is satisfied. Since the empty set contains no elements, the sentence "$x \in \varnothing$" is false for every element x. For this reason, you should avoid such notation, for example, when solving equations or inequalities.

Kind of the opposite of the empty set is the **universe** (otherwise known as the space), being the set that contains all the objects that we are interested in at the moment. The universe is denoted by the letter U. Usually the form of the universe results from the context. For every set A it is true that $A \subset U$.

Complement of a set A, denoted by A' (or A^C), is the set of all the elements not belonging to A. In other words, the complement of A may be described like this: $A' = \{x : x \notin A\}$ or like this: $x \in A' \Leftrightarrow \neg\, x \in A$. Graphical interpretation of the complement is shown in Fig. 2.1.

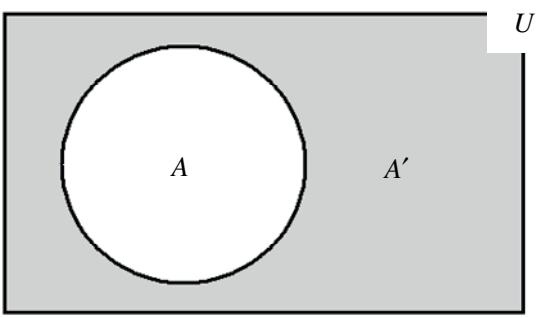


Fig. 2.1. Complement of a set

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Union of sets A and B, denoted by $A \cup B$, is the set consisting of all the elements that belong to at least one of the sets A and B. In other words it is the set of all the elements, which belong to A **or** to B. The union of A and B may be thus described like this: $A \cup B = \{x : x \in A \vee x \in B\}$ or like this: $x \in A \cup B \Leftrightarrow x \in A \vee x \in B$. The union is illustrated in Fig. 2.2.

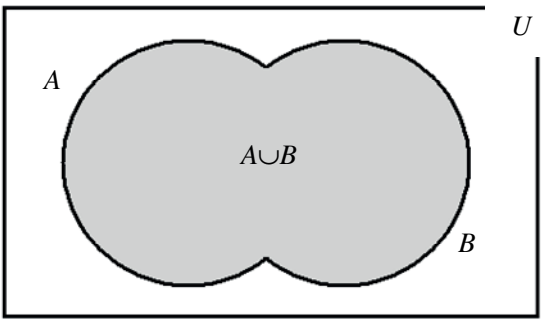


Fig. 2.2. Union of two sets

Intersection of sets A and B, denoted by $A \cap B$, is the set consisting of all the elements that belong to **both** A and B. In other words it is the set of all the elements, which belong to A **and** to B. Intersection of A and B may be thus described like this: $A \cap B = \{x : x \in A \wedge x \in B\}$ or like this: $x \in A \cap B \Leftrightarrow x \in A \wedge x \in B$. The intersection is illustrated in Fig. 2.3.

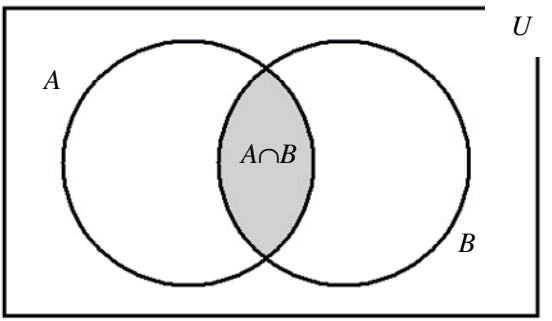


Fig. 2.3. Intersection of two sets

Difference of sets A and B, denoted by $A \setminus B$ is the set consisting of all the elements that belong **only** to A. In other words it is the set of all the elements

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

that belong to A **and** do not belong to B. The difference of A and B may be thus described like this: $A \setminus B = \{x : x \in A \wedge x \notin B\}$ or like this: $x \in A \setminus B \Leftrightarrow x \in A \wedge \neg\, x \in B$. The difference is illustrated in Fig. 2.4.

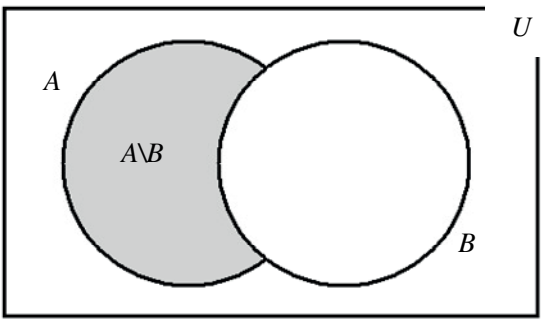


Fig. 2.4. Difference of two sets

Similarly as in the case of the logic formulae, also here we can list some important formulae which are always true, called the **set algebra laws**. They are listed in Table 2.1.

Table 2.1. Important set algebra laws

Law	Name
$(A')' = A$	Involution law
$A \cap A' = \varnothing$	Complement laws
$A \cup A' = U$	
$A \cup B = B \cup A$	Commutative laws
$A \cap B = B \cap A$	
$A \cup (B \cup C) = (A \cup B) \cup C$	Associative laws
$A \cap (B \cap C) = (A \cap B) \cap C$	
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributive laws
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	
$A \cap U = A$	Absorption laws
$A \cup U = U$	
$A \cap \varnothing = \varnothing$	
$A \cup \varnothing = A$	
$(A \cup B)' = A' \cap B'$	De Morgan's laws
$(A \cap B)' = A' \cup B'$	

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

As you will notice, these laws are analogous to the tautologies from the previous chapter (Table 1.3). Indeed, in order to prove them we will use respective tautologies.

We are mostly interested in the **sets of numbers** (or number sets), i.e. the sets having only the members being numbers. A number of such sets you can find already in the previous chapter.

Particularly interesting sets of numbers are the following infinite sets:

- the set of natural numbers: $N = \{0, 1, 2, 3, \ldots\}$;
- the set of integers: $Z = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$;
- the set of rational numbers: $Q = \{x : \exists p \in Z \, \exists \, q \in Z \setminus \{0\} : x = p/q\}$, i.e. the set of fractions having integer numerators and denominators, for example 1/2, –3/17, 43/3;
- the set of real numbers: R; formal definition of this set is rather tricky (and completely not useful for our purposes), so let us use some intuition and describe it as "the set of all the numbers that can be written using the decimal representation, finite or infinite"; the rational numbers have also finite or infinite decimal representations, but in the case of infinite representation it has a period (it means that starting from some point the same sequence of digits repeats infinitely many times); there are also real numbers which are not rational – their decimal representations are infinite and do not have periods, examples of such numbers are: $-\sqrt{2} = -1.41421...$, $e = 2.71818...$ and $\pi = 3.14159...$

The following inclusion relations are satisfied for the sets listed above: $N \subset Z \subset Q \subset R$. In other words, all the natural numbers are integers, all the integers (so also natural numbers) are rational numbers, all rational numbers (including natural numbers and integers) are real.

Using the given notation together with symbol "+" or "–", we can restrict ourselves to positive or negative numbers. For example "R_-" denotes the set of positive real numbers, and "$Z_+ \cup \{0\}$" is the set of positive integers together with the number 0, so in other words – the set of natural numbers.

The subsets of N, Z, Q and R may be described by enumeration or using the formulae (just as in the beginning of this section). The subsets of R may be also described as **intervals**. The intervals can be:

- **open**; for instance (1, 2.2) is the set of all the real numbers greater than 1 and lower than 2.2; (3.27, ∞) in turn is the set of all the real numbers greater than 3.27 (the symbol "∞" denotes infinity);
- **closed** (from one side or from both sides); for example left-closed interval [2.11, ∞) is the set of all the real numbers greater than or equal to 2.11; closed interval [4, 7] is in turn the set of all the real numbers greater than or equal to 4 and not greater than 7.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1
The following sets are given: $A = \{1, 2, 3\}$, $B = \{1, 2, 4\}$, $C = \{1, 2\}$ and $D = \varnothing$. Find their cardinalities.

Solution
The set A has three members: 1, 2 i 3. Its cardinality is then $|A| = 3$. Similarly $|B| = 3$ and $|C| = 2$. D is the empty set, so it has no members. Thus $|D| = 0$.

Example 2
Read the following notation and then list all the elements of the given set:

$$A = \{x \in N : x < 5\}.$$

Solution
"A is the set of the natural numbers x such that x is lower than 5". Its members are: 0, 1, 2, 3 and 4.

Example 3
Read the following notation and then list all the elements of the given set:

$$X = \{x : -10 < x < 10 \wedge (\exists y \in Z : x = 3y)\}.$$

Solution
"X is the set of x such that x greater than –10, x is lower than 10 and there exists an integer y such that x is equal to $3y$". Two first conditions are trivial and the third one means that x is divisible by 3. There are seven numbers being members of X: –9, –6, –3, 0, 3, 6 and 9.

Example 4
The following sets are given: $A = \{1, 2, 3, 4\}$, $B = (1, 4)$, $C = [1, 4]$, $D = N \cap C$, $E = \{x \in Z : x \in B\}$. List all the relations of inclusion and equality.

Solution
The set A has four members: 1, 2, 3 and 4. The sets B and C are intervals from 1 to 4, with the difference that the ends belong to C (closed interval), but not to B (open interval). The set D is intersection of C and the set of natural numbers N, so it has members 1, 2, 3 and 4. Similarly the set E consists of the numbers 2 and 3 (only these two numbers are members of B – 1 and 4 are the ends of the interval and B is open).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Taking into account the above considerations we conclude that the following relations occur (we skip the trivial relations of the form $X = X$ and $X \subset X$): $A = D$, $A \subset C$, $A \subset D$, $B \subset C$, $D \subset A$, $D \subset C$, $E \subset A$, $E \subset B$, $E \subset C$, $E \subset D$.

Example 5
The following sets are given: $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$. The universe is $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Find the complements, union, intersection and differences of A and B.

Solution
The complement of a set X is the set of all the elements not being members of X (but being members of the universe). Thus $A' = \{5, 6, 7, 8\}$ and $B' = \{1, 2, 7, 8\}$. The union consists of all the elements being in one set or another. Thus $A \cup B = \{1, 2, 3, 4, 5, 6\}$. Intersection consists of all the elements being members of both sets, so $A \cap B = \{3, 4\}$. The difference $A \setminus B$ consists of all the members of A not being members of B, so $A \setminus B = \{1, 2\}$. Similarly $B \setminus A = \{5, 6\}$.

Example 6
The following sets are given: $A = (1, 4]$ and $B = [3, 6]$. Find the complements, union, intersection and differences of A and B.

Solution
As the universe has not been defined, we take by default $U = R$. It follows that $A' = (-\infty, 1] \cup (4, \infty)$. Observe that the set A' has been written with the symbol of union, because it consists of two disjoint intervals. Another thing that you have to remember of is that we change the type of the interval: as A is left-open, A' has to be closed in the same point (as 1 is not a member of A, it must be a member of A'). Similarly with 4: A is closed, so A' is open. Continuing this way, we can find $B' = (-\infty, 3) \cup (6, \infty)$. Being careful about the ends of the intervals, we can also find the remaining sets: $A \cup B = (1, 6]$, $A \cap B = [3, 4]$, $A \setminus B = (1, 3)$ and $B \setminus A = (4, 6]$.

Example 7
Using the respective tautology prove the involution law.

Solution
Recall that the involution law has the form: $(A')' = A$. There is negation in the definition of complement. Thus the respective tautology is the double negation law: $\neg\neg p \Leftrightarrow p$ (Table 1.3 in the previous chapter).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Assume that some element x is a member of the set on the left side of the involution law. W can write it as follows: $x \in (A')'$. Using the definition of complement, we obtain an equivalent notation $\neg\neg\, x \in A$. Let us denote by p the statement "$x \in A$". This way we obtain the following notation: $\neg\neg\, p$. The double negation law says that this is equivalent to p, that is to "$x \in A$". Finally we obtain that $x \in (A')' \Leftrightarrow x \in A$. As we chose x arbitrarily, it means that $(A')' = A$.
In the remainder of this chapter the reasoning as above we will write in a shorter version, without comments. In the case of this example such a shorter form would look like:

$x \in (A')' \Leftrightarrow \neg\neg\, x \in A \Leftrightarrow$ (substitution of p for "$x \in A$")
$\Leftrightarrow \neg\neg\, p \Leftrightarrow$ (double negation law)
$\Leftrightarrow p \Leftrightarrow$ (back to the previous notation)
$\Leftrightarrow x \in A$.

Example 8
Using the respective tautology prove the first complement law.

Solution
The first complement law has the form: $A \cap A' = \varnothing$. The respective tautology is the 1st negation law: $p \wedge \neg\, p \Leftrightarrow 0$. The proof is as follows:

$x \in A \cap A' \Leftrightarrow x \in A \wedge x \in A' \Leftrightarrow x \in A \wedge \neg\, x \in A \Leftrightarrow$
(substitution of p for "$x \in A$")
$\Leftrightarrow p \wedge \neg\, p \Leftrightarrow$ (1st negation law)
$\Leftrightarrow 0$.

The false statement at the end means that also the initial statement was false. As we chose x arbitrarily, it means that no element is member of $A \cap A'$, thus $A \cap A'$ is the empty set ($A \cap A' = \varnothing$).

Example 9
Using the respective tautology prove the first distributive law.

Solution
The first distributive law has the form: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. The respective tautology is the first distributive law: $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$. The proof is as follows:

$x \in A \cup (B \cap C) \Leftrightarrow x \in A \vee x \in (B \cap C) \Leftrightarrow x \in A \vee (x \in B \wedge x \in C) \Leftrightarrow$
(substitution of p for "$x \in A$", q for "$x \in B$" and r for "$x \in C$")
$\Leftrightarrow p \vee (q \wedge r) \Leftrightarrow$ (first distributive law)

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$\Leftrightarrow (p \vee q) \wedge (p \vee r) \Leftrightarrow$ (back to the previous notation)
$\Leftrightarrow (x \in A \vee x \in B) \wedge (x \in A \vee x \in C) \Leftrightarrow$
$\Leftrightarrow x \in (A \cup B) \wedge x \in (A \cup C) \Leftrightarrow$
$\Leftrightarrow x \in (A \cup B) \cap (A \cup C)$.
This ends the proof.

EXERCISES

1. Find the cardinalities of the following sets:
   a) $A = \{1, 2, 3\}$;
   b) $B = \{1, 2, 3, \ldots, 10\}$;
   c) $C = \{3, 6, 9, \ldots, 27\}$;
   d) $D = \{12, 14, 16, \ldots, 26\}$.
2. Read the following definitions, and then list all the elements of the sets or write them in the form of intervals:
   a) $A = \{x \in Z : |x| < 3\}$;
   b) $A = \{x \in N : |x| < 3\}$;
   c) $A = \{x \in R : |x| < 3\}$;
   d) $A = \{x \in Z : x > 2 \wedge x \leq 5\}$;
   e) $A = \{x \in R : x > 2 \wedge x \leq 5\}$;
   f) $A = \{x \in N : x \leq 1 \vee (x \geq 7 \wedge x < 8)\}$;
   g) $A = \{x \in R : x \leq 1 \vee (x \geq 7 \wedge x < 8)\}$.
3. For given sets A and B find their complements, union, intersection and both differences. If the universe is not given, take by default $U = R$.
   a) $A = \{1, 2\}$, $B = \{2, 4\}$, $U = \{1, 2, 3, 4, 5\}$;
   b) $A = \{1, 3, 5, 7\}$, $B = \{2, 4, 5, 6\}$,
      $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$;
   c) $A = \{x \in N : x < 5\}$, $B = \{x \in N : 2 < x < 8\}$,
      $U = \{x \in N : x < 9\}$;
   d) $A = \{x \in Z : -2 \leq x \leq 5\}$, $B = \{x \in Z : -4 < x < 3\}$, $U = A \cup B$;
   e) $A = (2, 5]$, $B = [3, 4]$;
   f) $A = (-2, 6]$, $B = [-3, \infty)$;
   g) $A = (-\infty, 0)$, $B = (0, \infty)$.
4. List all the relations of inclusion and equality for the following families of sets:
   a) $A = \{1, 2, 5\}$, $B = \{2, 5\}$, $C = \{2, 3, 5\}$, $D = \{1, 2\}$, $E = \{3, 5\}$;
   b) $A = \{1, 2, 5\}$, $B = \{x \in N : x < 5\}$, $C = \{x \in N : x < 6\}$,
      $D = \{x \in N : x \leq 5\}$, $E = \{x \in N : x \leq 6 \vee x \geq 10\}$;
   c) $A = \{-2, -1, 2, 4, 8\}$, $B = \{x \in Z : x \geq -3 \wedge x \leq 8\}$, $C = (-3, 8)$,
      $D = [-3, 8]$, $E = \{x \in Z : x > -3 \wedge x < 8\}$;
   d) $A = (-\infty, 2]$, $B = (-\infty, 2)$, $C = (-5, 2]$, $D = [-5, 2)$, $E = (-5, 2)$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

e) $A = \{x \in R : x < 5\}$, $B = \{x \in R : x \leq 5\}$, $C = (-\infty, 5]$, $D = (-\infty, 5)$, $E = [-5, 5]$.

5. For each of the following set algebra laws find the respective tautology, and then use it to prove the law:
   a) second complement law: $A \cup A' = U$;
   b) first commutative law: $A \cup B = B \cup A$;
   c) second commutative law: $A \cap B = B \cap A$;
   d) first associative law: $A \cup (B \cup C) = (A \cup B) \cup C$;
   e) second associative law: $A \cap (B \cap C) = (A \cap B) \cap C$;
   f) second distributive law: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
   g) first absorption law: $A \cap U = A$;
   h) second absorption law: $A \cup U = U$;
   i) third absorption law: $A \cap \varnothing = \varnothing$;
   j) fourth absorption law: $A \cup \varnothing = A$;
   k) first De Morgan's law: $(A \cup B)' = A' \cap B'$;
   l) second De Morgan's law: $(A \cap B)' = A' \cup B'$.

SOLUTIONS

1. a) $|A| = 3$; b) $|B| = 10$; c) $|C| = 9$; d) $|D| = 8$.
2. Sample answers:
   a) A is the set of integers x such that the absolute value of x is lower than 3; $A = \{-2, -1, 0, 1, 2\}$;
   b) A is the set of natural numbers x such that the absolute value of x is lower than 3; $A = \{0, 1, 2\}$;
   c) A is the set of real numbers x such that the absolute value of x is lower than 3; $A = (-3, 3)$;
   d) A is the set of integers x such that x is greater than 2 and x is lower than or equal to 5; $A = \{3, 4, 5\}$;
   e) A is the set of real numbers x such that x is greater than 2 and x is lower than or equal to 5; $A = (2, 5]$;
   f) A is the set of natural numbers x such that x is lower than or equal to 1 or x is greater than or equal to 7 and lower than 8; $A = \{0, 1, 7\}$;
   g) A is the set of natural numbers x such that x is lower than or equal to 1 or x is greater than or equal to 7 and lower than 8; $A = (-\infty, 1] \cup [7, 8)$.
3. Desired sets:
   a) $A' = \{3, 4, 5\}$, $B' = \{1, 3, 5\}$, $A \cup B = \{1, 2, 4\}$, $A \cap B = \{2\}$, $A \setminus B = \{1\}$, $B \setminus A = \{4\}$;
   b) $A' = \{2, 4, 6, 8, 9, 10\}$, $B' = \{1, 3, 7, 8, 9, 10\}$, $A \cup B = \{1, 2, 3, 4, 5, 6, 7\}$, $A \cap B = \{5\}$, $A \setminus B = \{1, 3, 7\}$, $B \setminus A = \{2, 4, 6\}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) $A' = \{5, 6, 7, 8\}$, $B' = \{0, 1, 2, 8\}$,
$A \cup B = \{0, 1, 2, 3, 4, 5, 6, 7\}$, $A \cap B = \{3, 4\}$, $A \setminus B = \{0, 1, 2\}$,
$B \setminus A = \{5, 6, 7\}$;

d) $A' = \{-3\}$, $B' = \{3, 4, 5\}$, $A \cup B = \{-3, -2, -1, 0, 1, 2, 3, 4, 5\}$,
$A \cap B = \{-2, -1, 0, 1, 2\}$, $A \setminus B = \{3, 4, 5\}$, $B \setminus A = \{-3\}$;

e) $A' = (-\infty, 2] \cup (5, \infty)$, $B' = (-\infty, 3) \cup (4, \infty)$, $A \cup B = (2, 5]$,
$A \cap B = [3, 4]$, $A \setminus B = (2, 3) \cup (4, 5]$, $B \setminus A = \emptyset$;

f) $A' = (-\infty, -2] \cup (6, \infty)$, $B' = (-\infty, -3)$, $A \cup B = [-3, \infty)$,
$A \cap B = (-2, 6]$, $A \setminus B = \emptyset$, $B \setminus A = [-3, -2] \cup (6, \infty)$;

g) $A' = [0, \infty)$, $B' = (-\infty, 0]$, $A \cup B = (-\infty, 0) \cup (0, \infty) = R \setminus \{0\}$,
$A \cap B = \emptyset$, $A \setminus B = (-\infty, 0)$, $B \setminus A = (0, \infty)$.

4. In the answers, the relations of a set with itself have been skipped:
   a) $B \subset A$, $B \subset C$, $D \subset A$, $E \subset C$;
   b) $A \subset C$, $A \subset D$, $A \subset E$, $B \subset C$, $B \subset D$, $B \subset E$, $C \subset D$, $C = D$, $C \subset E$, $D \subset C$, $D \subset E$;
   c) $A \subset B$, $A \subset D$, $B \subset D$, $C \subset D$, $E \subset B$, $E \subset C$, $E \subset D$;
   d) $B \subset A$, $C \subset A$, $D \subset A$, $D \subset B$, $E \subset A$, $E \subset B$, $E \subset C$, $E \subset D$;
   e) $A \subset B$, $A \subset C$, $A \subset D$, $A = D$, $B \subset C$, $B = C$, $C \subset B$, $D \subset A$, $D \subset B$, $D \subset C$, $E \subset B$, $E \subset C$.
5. In the answers the appropriate tautologies have been indicated, sometimes followed by additional hints:
   a) excluded middle law: $p \vee \neg p \Leftrightarrow 1$; arbitrary element x belongs to a set if and only if it is the universe;
   b) first commutative law: $p \vee q \Leftrightarrow q \vee p$;
   c) second commutative law: $p \wedge q \Leftrightarrow q \wedge p$;
   d) first associative law: $p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$;
   e) second associative law: $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$;
   f) second distributive law: $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$;
   g) first absorption law: $p \wedge 1 \Leftrightarrow p$; arbitrary element x belongs to a set if and only if it is the universe;
   h) second absorption law: $p \vee 1 \Leftrightarrow 1$; arbitrary element x belongs to a set if and only if it is the universe;
   i) third absorption law: $p \wedge 0 \Leftrightarrow 0$; no element x belongs to a set if and only if it is the empty set;
   j) fourth absorption law: $p \vee 0 \Leftrightarrow p$; no element x belongs to a set if and only if it is the empty set;
   k) first De Morgan's law: $\neg (p \vee q) \Leftrightarrow (\neg q \wedge \neg p)$;
   l) second De Morgan's law: $\neg (p \wedge q) \Leftrightarrow (\neg q \vee \neg p)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2.2. CARTESIAN PRODUCT

THEORY IN A NUTSHELL

The Cartesian product $A \times B$ of two sets A and B is the set of all the ordered pairs (x, y) such that $x \in A$ and $y \in B$. It can be written as:

$$A \times B = \{(x, y) : x \in A \wedge y \in B\}.$$

Remember that in the Cartesian product the order of the elements matters, so $A \times B$ is not the same as $B \times A$.
The notion of Cartesian product can be generalized to a larger number (three, four and so on) of sets – in this case, instead of ordered pairs we consider longer sorted sequences (respectively, triple, quadruple and so on).
If we consider the Cartesian product of A with itself, instead of $A \times A$ we simply write A^2. This notation is often used in the case of the sets of numbers: for example, R^2 is the set of pairs of real numbers, that is, the set of coordinates of points on the plane. Similarly, R^3 is the set of coordinates of points in three–dimensional space. This kind of notation is called the **power of a set**.

EXAMPLES

Example 1
Present the Cartesian product of the following set by listing all its members: $A = \{1, 2, 5\}$ and $B = \{2, 3, 4\}$.

Solution
We have to list all the ordered pairs such that first element is member of A and the second one is member of B. One of the members of A is 1 (note that we don't say that 1 is first member – remember that the order does not matter in the case of a set, although it does in the case of an ordered pair or a sequence). Now we join 1 with all the members of B. This way we obtain the ordered pairs (1, 2), (1, 3) and (1, 4). Recall once again – the pairs may be listed in any order (for example (1, 3), (1, 4) and (1, 2)), but the elements of each pair have predefined order (thus for example the ordered pair (2, 1) does not belong to the Cartesian product $A \times B$, because is not the same as the ordered pair (1, 2)).
Proceeding similarly with other members of A (i.e. 2 and 5), we obtain six other ordered pairs (for each of two members of A we have three distinct members of B). Finally we obtain the following result:
$A \times B = \{(1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (5, 2), (5, 3), (5, 4)\}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 2
Present the graphical interpretation (on the plane) of the Cartesian product of the sets $A = (–1, 5]$ and $B = (2, 4]$.

Solution
We have to plot all the points on the graph whose first coordinate belongs to the interval A, and the second to the interval B. The easiest way you can do it is the following: mark the ends of A on the horizontal axis $0x_1$, and the ends of B on the vertical axis $0x_2$ and then draw the dashed lines containing these points, orthogonal to the axes, as in Fig. 2.5a. Then fill the area between those lines and bold the lines corresponding to the ends of the intervals that belong to them, as in Fig. 2.5b.

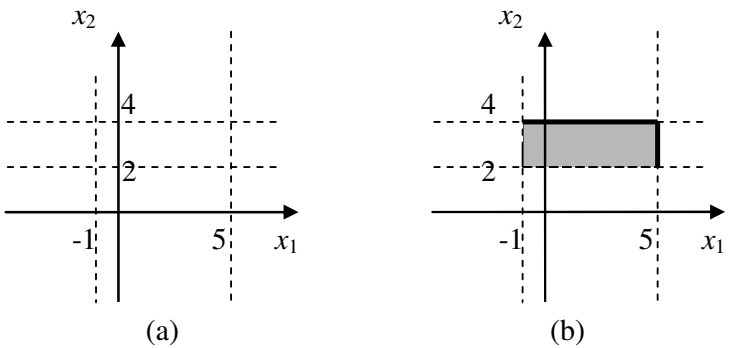


Fig. 2.5. Cartesian product

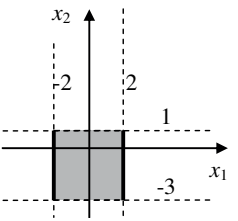
EXERCISES

1. List all the elements of the Cartesian products $A \times B$ and $B \times A$:
   a) $A = \{3, 5, 7\}$, $B = \{4, 6\}$;
   b) $A = \{x \in N : x \leq 2\}$, $B = \{x \in Z : |x| < 2\}$;
   c) $A = \{x \in N : x \geq 3 \wedge x \leq 5\}$, $B = \{x \in Z : x \geq -2 \wedge x \leq 0\}$.
2. Present the graphical image of the Cartesian products $A \times B$ and $B \times A$:
   a) $A = [–2, 2]$, $B = (–3, 1)$; b) $A = (–\infty, 5)$, $B = [2, \infty)$;
   c) $A = (–3, 2) \cup [2, \infty)$, $B = [2, 4]$; d) $A = R_+$, $B = R_+ \cup \{0\}$.
3. Write the following powers of sets in the form of products and find two sample elements:
   a) A^4, where $A = \{1, 4, 11\}$; b) R^5;
   c) N^3; d) Z^4.

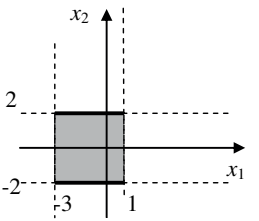
Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. Desired sets:
   a) $A \times B = \{(3, 4), (3, 6), (5, 4), (5, 6), (7, 4), (7, 6)\}$,
   $B \times A = \{(4, 3), (4, 5), (4, 7), (6, 3), (6, 5), (6, 7)\}$;
   b) $A \times B = \{(0, -1), (0, 0), (0, 1), (1, -1), (1, 0), (1, 1), (2, -1), (2, 0), (2, 1)\}$,
   $B \times A = \{(-1, 0), (-1, 1), (-1, 2), (0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2)\}$;
   c) $A \times B = \{(3, -2), (3, -1), (3, 0), (4, -2), (4, -1), (4, 0), (5, -2), (5, -1), (5, 0)\}$, $B \times A = \{(-2, 3), (-2, 4), (-2, 5), (-1, 3), (-1, 4), (-1, 5), (0, 3), (0, 4), (0, 5)\}$.
2. Solutions on the pictures below:
   a)

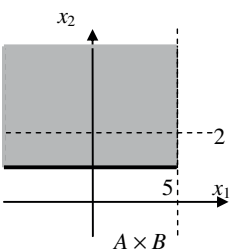


$A \times B$

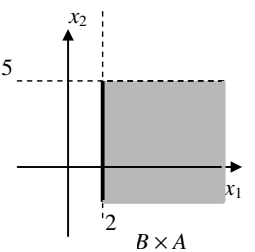


$B \times A$

   b)

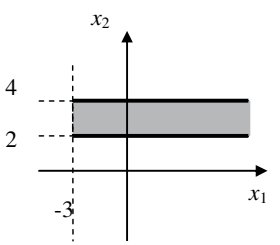


$A \times B$

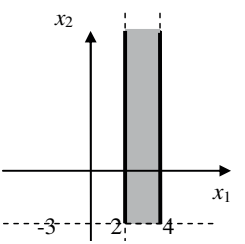


$B \times A$

   c)



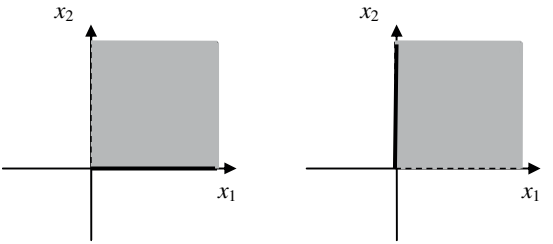
$A \times B$



$B \times A$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

d)



$A \times B$ $B \times A$

3. Products and sample elements:
   a) $A \times A \times A \times A$; (1, 1, 4, 4); (11, 11, 11, 1);
   b) $R \times R \times R \times R \times R$; (2, 3, –1.4, –2.237, π); (1, –1, 1, 0, 0);
   c) $N \times N \times N$; (0, 0, 2); (153, 11, 1);
   d) $Z \times Z \times Z \times Z$; (–2, 0, 5, 1000); (–33, –2, –1, 666).

2.3. THE INCLUSION–EXCLUSION PRINCIPLE

THEORY IN A NUTSHELL

The sum of two sets A and B is the set of all those elements that belong **to *A* or to *B***. Let us try to determine the cardinality of $A \cup B$.
We cannot simply add the cardinalities of A and B, because the elements belonging to both sets would then be counted twice. Hence the conclusion that the number of such elements counted twice must be subtracted. This leads us to the following formula:

$|A \cup B| = |A| + |B| - |A \cap B|$.

It can be generalized to a larger number of sets. The formula is given in the Theorem 2.1.

Theorem 2.1. The Inclusion–Exclusion Principle
Let the sets $A_1, A_2, \ldots, A_n$ be given. Then the following holds:

$$\left|A_1 \cup A_2 \cup \ldots \cup A_n\right| =$$
$$= \sum_{1 \le i \le n} \left|A_i\right| - \sum_{1 \le i < j \le n} \left|A_i \cap A_j\right| + \sum_{1 \le i < j < k \le n} \left|A_i \cap A_j \cap A_k\right| - \ldots \pm \left|A_1 \cap A_2 \cap \ldots \cap A_n\right|.$$

The symbol Σ (sigma) denotes the sum. More about this symbol can be found in Chapter 7. Now it is important that the first sum is the sum of the cardinalities of all the sets A_i, the second one is the sum of the cardinalities of all the

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

intersections of two distinct sets A_i and A_j and so on. The meaning of the formula is as follows. If you want to count the members of a union of n sets, sum up the cardinalities of all those sets, subtract the cardinalities of all the intersections of two distinct sets, add the cardinalities of all the intersections of three distinct sets, subtract the cardinalities of all the intersections of four distinct sets… and so on, alternately add and subtract until you reach the intersection of all the n sets.

EXAMPLES

Example 1
Two sets A and B are given. We know that $|A| = 5$, $|B| = 7$ and $|A \cap B| = 3$. Find $|A \cup B|$.

Solution
$|A \cup B| = |A| + |B| - |A \cap B| = 5 + 7 - 3 = 9$.

Example 2
How many cards being red or being honors there are in an ordinary set of 52 cards?

Solution
Assume that A is the set of honors. There are 16 of them (4 Jacks, 4 Queens, 4 Kings, 4 Aces). Let B be the set of red cards. There are 26 of them (13 hearts and 13 diamonds). Moreover there are 8 red honors (2 Jacks, 2 Queens, 2 Kings, 2 Aces). We have then $|A| = 16$, $|B| = 26$ and $|A \cap B| = 8$. Hence $|A| + |B| - |A \cap B| = 26 + 16 - 8 = 34$. It means that there are 34 cards being red or being honors.

Example 3
At the Faculty X on the first year of studying one hundred and eighty people. One hundred of them have the second term exam in mathematics, one in microeconomics and one in computer science. It is known, moreover, that eighty people have the second term exam in mathematics and microeconomics, seventy in mathematics and computer science, seventy-five in the computer science and microeconomics, and sixty in all three subjects. How many people passed all three subjects in the first term?

Solution
Let A be the set of the students that have to take the second term exam in mathematics, B – the set of the students that have to take the second term exam in microeconomics and C – the set of the students that have to take the second

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

term exam in computer science. It follows that $|A| = |B| = |C| = 100$, $|A \cap B| = 80$, $|A \cap C| = 70$, $|B \cap C| = 75$, $|A \cap B \cap C| = 60$. By applying the Inclusion–Exclusion Principle we obtain:

$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| =$
$= 100 + 100 + 100 - 80 - 70 - 75 + 60 = 135.$

This means that one hundred thirty five people must take at least one second term exam. This in turn means that the remaining forty-five people do not have to take any second term exam.

EXERCISES

1. University consists of two faculties: X and Y. There are 2500 students at the faculty X, 2,000 at the faculty Y, 500 of which study at both faculties. How many students there are in total at the University?
2. A beer producer performed market research in a city of 100,000 habitants. It came out that 20,000 habitants drink beer A, 10,000 beer B, 40,000 beer C. Moreover half of those who drink B, drink also A, 4,000 drink B and C, and A and C is the favorite beer of 8,000 people, among which 3,000 drink also B. How many habitants do not drink beer at all?
3. University consists of two faculties: X and Y. There are 1,500 students at the faculty X, and 2,000 at Y. It is also known that there are exactly 2,500 students at the University. How many students are members of both faculties?
4. Transportation company owns 85 trucks; 20 of them have problems with the exhaust system, 30 have damaged shocks, and in 10 of them the brakes don't work; 10 trucks have damaged shocks and exhaust system, 7 damaged brakes and shocks, and 7 damaged brakes and exhaust system; in 5 trucks all three systems don't work properly. Is the owner of the company able to use at least half of the trucks today?
5. Employees of a financial company can obtain a quarterly bonus if they fulfil at least one of the following conditions: (1) selling the mortgages of the total value at least 1,000,000 PLN, (2) selling the cash loans of the total value at least 100,000 PLN, (3) selling the life insurances of the total value at least 1,000,000 PLN. It is known that seven employees fulfilled condition (1) (among them three also condition (2) and two also condition (3)), five employees condition (2) (two of them also condition (3)) and three employees condition (3). Only one employee fulfilled all three conditions. How many employees will obtain the bonus?

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. $2{,}500 + 2{,}000 - 500 = 4{,}000$.
2. The number of people drinking beer:
   $20{,}000 + 10{,}000 + 40{,}000 - 5{,}000 - 4{,}000 - 8{,}000 + 3{,}000 = 56{,}000$.
   The number of people not drinking beer: $100{,}000 - 56{,}000 = 44{,}000$.
3. Let x be the missing value. $1{,}500 + 2{,}000 - x = 2{,}500$, thus $x = 1{,}000$.
4. Number of damaged cars: $20 + 30 + 10 - 10 - 7 - 7 + 5 = 41$. Number of working cars: $85 - 41 = 44 > 0{,}5 \cdot 85$. Yes, he can.
5. $7 + 5 + 3 - 3 - 2 - 2 + 1 = 9$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 3
RELATIONS

3.1. DESCRIBING THE RELATIONS

THEORY IN A NUTSHELL

Intuitively, relation can be interpreted as a link between the two sets of elements, and not all of those elements must be linked. The order of elements matters ("x is related to y" means something other than "y is related to x"). Does it remind you something?

Definition 3.1. Relation
Relation is a subset of the Cartesian product of two sets.

Definition 3.2. Domain and counterdomain
Assume that there is a relation $P \subset A \times B$. Then the set A is called the **domain** of P and denoted by $D(P)$, while B is called the **counterdomain** of P and denoted by $D^{-1}(P)$.

If $x \in A$ and $y \in B$, then the fact that "x is P-related to y" is denoted by "xPy".
Relations can be described in several ways.
If the sets A and B are finite, you can list all the pairs of elements that make up a relation, or present it in the form of a **matrix**. The matrix of relation is a rectangular array filled with numbers 0 or 1. Her rows represent the members of the domain, and columns – the members of the counterdomain. If element is related to another, there is 1 at the intersection of the corresponding row and column of otherwise there is 0. In the case of finite sets, a relation can also be described using a **graph**. In this case we write all the elements of domain and counterdomain and connect with arcs (arrows) those that are related. If $A = B$ (i.e., domain and counterdomain are the same set), it is convenient to write all the elements of domain (and counterdomain) only once.
Each relation can also be described by a formula and presented on a plot. In this case domain and counterdomain do not have to be finite. In this chapter we will deal only with finite relations, however, you will use plots later in the book, when we discuss a particular type of relations, which are functions. Most of the concepts in this chapter will be presented on the example of matrices and graphs.
The relation is a set of ordered pairs, and so one can perform on them the same actions as on sets (union, intersection, difference, complement). You only have to remember two things: firstly, the two relations must have the same domain and the same counterdomain, and secondly, universe is the Cartesian product of the domain and counterdomain.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

We will be particularly interested in two new actions referred specifically to the relations:

- **Composition of relations**. If $P \subset A \times B$ and $R \subset B \times C$, then the composition $R \circ P$ (watch out for the order!) consists of the pairs (x, y) such that there is z, for which the following holds: xPz and zRy. More formally:

  $$xR \circ Py \Leftrightarrow \exists z : xPz \wedge zRy .$$

- **Inverse relation**. The inverse relation P^{-1} consists of the pairs (x, y) such that yPx. More formally:

  $$xP^{-1}y \Leftrightarrow yPx .$$

EXAMPLES

Example 1
Let $A = \{1, 2, 3\}$ and $B = \{0, 1, 2, 3\}$. Let $P \subset A \times B : xPy \Leftrightarrow x \geq y$. Write P using list, matrix and graph.

Solution
To list all members of the relation, we will check all the members of the Cartesian product $A \times B$, and choose only the related ones. Let us start with the pairs having 1 as the first element. There are four such pairs: (1, 0), (1, 1), (1, 2) and (1, 3). We have to check where the first element (i.e., x) is greater than or equal to the second one (i.e., y). This condition is satisfied in the case of pairs (1, 0) and (1, 1), so they are members of the relation P. The pairs (1, 2) and (1, 3) do not satisfy the condition, so they are not members of the relation. By checking pairs beginning with 2, we find that the elements of pairs (2, 0), (2, 1) and (2, 2) are related, which is not true in the case of pair (2, 3). The relation also includes all pairs starting with 3. Finally, it can be written as:

$P = \{(1, 0), (1, 1), (2, 0), (2, 1), (2, 2), (3, 0), (3, 1), (3, 2), (3, 3)\}$.

Now we are going to write the relation in the form of matrix. The domain A and the counterdomain is B, so the rows will be signed with numbers 1, 2 and 3, and columns – with the numbers 0, 1, 2 and 3. Let us begin with the first row. Remind how we listed the pairs being members of the relation. We started with the pairs where the first element was 1. So let's start now with the row signed "1". Since (as previously established), two of the mentioned pairs belong to the relation (pair (1, 0) and pair (1, 1)), we put 1 in the appropriate columns (signed

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

"0" and "1"). Other elements of the row (columns "2" and "3") will be filled with 0, since the two other couples do not belong to the relation. The operation result is shown in the matrix (a). Similarly, considering the row "2", we insert 1 only to the first three columns, because only pairs (2, 0), (2, 1) and (2, 2) are members of the relation (matrix (b)). In the third row there are only 1, because all pairs beginning with "3" belong to the relation (matrix (c)).

(a)

	0	1	2	3
1	1	1	0	0
2				
3				

(b)

	0	1	2	3
1	1	1	0	0
2	1	1	1	0
3				

(c)

	0	1	2	3
1	1	1	0	0
2	1	1	1	0
3	1	1	1	1

It remains to present the relation as a graph. To this end, we write all the members of the sets A and B, and connect the related elements with the arcs (arrows):

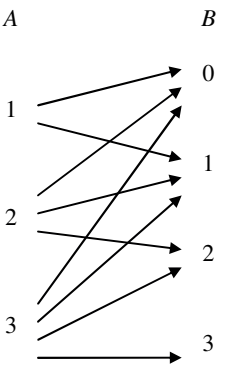


Fig. 3.1. Graph of relation

Example 2
Let $P \subset A^2 : xPy \Leftrightarrow x + y \geq 5$, where $A = \{1, 2, 3, 4\}$. List all the members of P, present its matrix and graph.

Solution
The relation is formed by all pairs of numbers whose sum is not less than 5, which can be written in the form:

$P = \{(1, 4), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4), (4, 1), (4, 2), (4, 3), (4, 4)\}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

To present the relation with the matrix, you have to note that in the first row (signed "1") we need to insert the number 1 in the column "4" (among all pairs starting from 1, only (1, 4) belongs to the relation), in the second row 1 must appear in the columns "3" and "4" (among all pairs starting with 2 only (2, 3) and (2, 4) belong to the relation), similarly in the third row 1 will be in all columns except "1" and the row "4" will be filled entirely with 1:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

As the set $A = \{1, 2, 3, 4\}$ is the domain, as well as the counterdomain of P, in order to present the relation it is enough to use one copy of each member of A, as in Fig. 3.2.

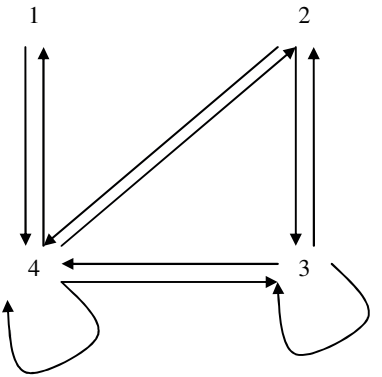


Fig. 3.2. Graph of relation

Pay attention to the pair (3, 3), represented by the arrow running from "3" to "3". Such arrows, beginning and ending at the same point (vertex), are called **loops**.

Example 3

Let $P \subset A^2 : xPy \Leftrightarrow x \geq y$ and let $Q \subset A^2 : xQy \Leftrightarrow x + y \geq 4$, where $A = \{1, 2, 3, 4\}$. Find both complements, union and intersection of P and Q, both differences and inverse relations. Present the results using matrices.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
The matrices of P and Q are as follows:

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Complement of a relation, as well as complement of each other set (the relation is a set of pairs!), is formed by these elements (i.e., those pairs), which do not belong to the relation. Let's start with the relation P. 1 is P-related to 1, and it is not P-related to 2, 3 nor 4. It follows that 1 is not P'-related to 1, but it is P'-related to 2, 3 and 4. For this reason, in the matrix of P' in row "1" the first element is 0, and the others are 1. We proceed this way with all remaining rows. In other words, to obtain the matrix of the relation P', all 0 in the matrix of P must be replaced with 1 and vice versa. Similarly we construct the matrix of Q'. The effect of these actions is shown below:

$$P' = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \qquad Q' = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

The union of two relations is the set of pairs which belong to at least one of them. For example, 1 is P-related to 1, Q-related to 3 and 4, so it follows that it is $(P \cup Q)$-related to 1, 3 and 4. If you use matrices to find the union, you have to rewrite all the ones, which occur in at least one of the input matrices. The intersection of two relations is the set of pairs that belong to both of them. If you use matrices to find the intersection, you have to rewrite all the ones, which occur in both input matrices. The union and intersection of the relations P and Q are shown below:

$$P \cup Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad P \cap Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The difference of two relations is formed by the pairs, which are in the first relation, but not in the other. Consider, for example, the difference $P \setminus Q$, for example the pairs starting with 2. As we can see, among the members of P there are only the pairs (2, 1) and (2, 2) and in the case of Q – the pairs (2, 2), (2, 3) and (2, 4). Of these, only the pair (2, 1) belongs to P and not to Q. If we refer to the matrix representation, in row "2" only the 1 in column "1" is present in the matrix of P, and not in the matrix of Q. Similarly, in row "2" only 1 in the columns "3" and "4" are present in the matrix of Q and not in P. Analyzing in a similar way the other rows, we obtain the matrices of differences:

$$P \setminus Q \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \qquad Q \setminus P \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix}$$

$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

To create the matrix of inverse relation, we rewrite the rows of the matrix as columns (or columns as rows)[4]. For example, the first row of the matrix of P consists of elements 1, 0, 0, and 0, so in the same order they form a first column of the matrix of P^{-1}. We apply the same operation to the other rows of the matrices of relations P and Q, which results with the matrices of the inverse relations. Please note that the relations Q and Q^{-1} are equal.

$$P^{-1} \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \qquad Q^{-1} \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix}$$

$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Example 4

Let $P \subset B \times C : xPy \Leftrightarrow x + y \leq 5$ and $Q \subset A \times B : xQy \Leftrightarrow x \cdot y \geq 5$, where $A = \{1, 2, 3\}$, $B = \{2, 3, 4, 5\}$ and $C = \{0, 2, 4\}$. Write them in the form of matrices and find the composition $P \circ Q$.

[4] This operation is called **transposition**. You will find more about transposition in the next chapter.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
The matrices look as follows:

$$P = \begin{array}{c} \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 2 & 4 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array} \qquad Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

By the definition of composition, it is a subset of $A \times C$ in this case. It consists of all the pairs (x, y) such that there is at least one $z \in B$ such that xQz and zPy. Verifying all the members of $A \times C$ may however take too much time, so we will use another method. We are looking for the composition $P \circ Q$, so for each element x being member of $D(Q)$ (i.e., of the domain of Q, in this case being equal to A) we check to which members of $D^{-1}(Q)$ (counterdomain of Q, i.e., B) it is Q-related. For each such element we check in turn, to which elements of the counterdomain of P (i.e., C), it is P-related. And those are the elements to which x is $(P \circ Q)$-related. The phases of determining the matrix of $P \circ Q$ are presented below.

$$P \circ Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 2 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ & & \\ & & \end{bmatrix} \end{array} \qquad P \circ Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 2 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ & & \end{bmatrix} \end{array} \qquad P \circ Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 2 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

Let us consider all the members of $D(Q)$. Let us start with 1. It is Q-related only to 5 (in the matrix of Q in the row "1" the only 1 is in the column "5"). Thus we find the row "5" in the matrix of P and we copy each 1 that we find. There is only one such 1, in the column "0". We copy it to the matrix of the composition, filling the rest of the row with 0 (left matrix above). We step to the next member of $D(Q)$, i.e., 2. In the row "2" of the matrix of Q we find 1 in the columns "3", "4" and "5", so we check the rows "3", "4" and "5" in the matrix of P. There are two 1 in the row "3": in the columns "0" and "2". In the row "4" there is only one 1 in the column "0". In the row "5" there is also only one 1 in the column "0". Finally, we have 1 in the columns "0" and "2" (it does not matter that in the column "0", 1 appears three times). For that reason, we insert 1 in the matrix of the composition only to the columns "0" and "2", and the remaining elements of the row "2" we fill with 0 (actually, we put only one 0 in the column "4"). The effect of our actions can be seen in the central matrix above. The last member of $D(Q)$ is 3. It is Q-related to 2, 3, 4 and 5. We check respective rows in the matrix

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

of P and we find out that taking them all into account we can insert 1 only to the columns "0" and "2" (both copies of 1 in the rows "2" and "3", moreover one 1 in the column "0" in the rows "4" and "5" of the matrix of P). Thus in the row "3" of the matrix of the composition there will be 1 in the columns "0" and "2". We see the final result in the last matrix above.

Example 5
Let $P \subset A^2 : xPy \Leftrightarrow x \geq y$ and $Q \subset A^2 : xQy \Leftrightarrow x + y \geq 4$, where $A = \{1, 2, 3, 4\}$. Find the compositions $P \circ Q$, $Q \circ P$, $P \circ P$ and $Q \circ Q$.

Solution
The matrices of P and Q are as follows:

$$\begin{array}{c|cccc} P & 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 \\ 4 & 1 & 1 & 1 & 1 \end{array} \qquad \begin{array}{c|cccc} Q & 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

In each of these cases we find the composition the same way as in the previous example. Note only that the two relations are defined on a product of A with itself, so also each of the compositions be a subset of A^2.
In the case of a $P \circ Q$ we check first which elements are Q-related to which, and then which are P-related to which. For example, let us consider 1. It is Q-related to 3 and 4. We move to the matrix of P, and we find that in the rows "3" and "4" 1 can be found in each column. So finally 1 is $(P \circ Q)$-related to all the elements of A. Similarly we proceed with the composition $Q \circ P$, but here we first check the relation P, then Q. For example, 1 is P-related only to 1, which is Q-related to 3 and 4. This means that 1 is $P \circ Q$-related to 3 and 4. The final effect has been presented below:

$$\begin{array}{c|cccc} P \circ Q & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 & 1 \end{array} \qquad \begin{array}{c|cccc} Q \circ P & 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

The compositions $P \circ P$ (i.e., P^2) and $Q \circ Q$ (i.e., Q^2) we find similarly, analyzing twice the matrix of P (or Q, respectively):

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$P^2 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 \\ 4 & 1 & 1 & 1 & 1 \end{array} \qquad Q^2 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

EXERCISES

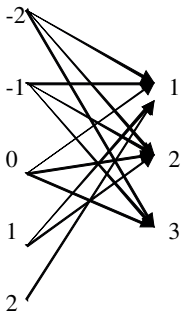
1. Present the following relations in the form of a list of pairs, a matrix and a graph (use the simplified version of graph where it is possible).
   a) $P \subset A \times B : xPy \Leftrightarrow x + y \leq 3$, $A = \{-2, -1, 0, 1, 2\}$, $B = \{1, 2, 3\}$;
   b) $P \subset A^2 : xPy \Leftrightarrow \exists\, k \in Z : x - y = 2k$, $A = \{-2, -1, 0, 1, 2\}$;
   c) $P \subset A \times B : \quad xPy \Leftrightarrow \exists\, k \in Z : x - y = 3k, \quad A = \{x \in N : x < 2\}$, $B = \{x \in N : x < 4\}$;
   d) $P \subset A^2 : xPy \Leftrightarrow x^2 < y^2$, $A = \{x \in Z : -1 \leq x \leq 3\}$.
2. Present the matrices of the following relations and then find their complements, unions, intersections, differences and inverse relations (where it is possible):
   a) $P \subset A \times B : xPy \Leftrightarrow x + y \leq 2$, $Q \subset A \times B : xQy \Leftrightarrow x - y \geq -1$, $A = \{-2, -1, 0, 1, 2\}$, $B = \{1, 2, 3\}$;
   b) $P \subset A^2 : xPy \Leftrightarrow \exists\, k \in Z : x - y = 2k$, $Q \subset A^2 : xQy \Leftrightarrow xy > 0$, $A = \{-1, 0, 1, 2\}$;
   c) $P \subset A \times B : xPy \Leftrightarrow xy \leq 2$, $Q \subset B \times C : xQy \Leftrightarrow xy \geq 1$, $A = \{-1, 0, 1, 2\}$, $B = \{1, 2, 3\}$, $C = \{x \in N : x < 2\}$.
3. Find all the existing compositions among the following: $P \circ Q$, $Q \circ P$, P^2 and Q^2:
   a) $P \subset A \times B : xPy \Leftrightarrow x + y \leq 2, \qquad Q \subset A \times B : xQy \Leftrightarrow x - y \geq -1$, $A = \{-2, -1, 0, 1, 2\}$, $B = \{1, 2, 3\}$;
   b) $P \subset A^2 : xPy \Leftrightarrow \exists\, k \in Z : x - y = 2k$, $Q \subset A^2 : xQy \Leftrightarrow xy > 0$, $A = \{-1, 0, 1, 2\}$;
   c) $P \subset A \times B : xPy \Leftrightarrow xy \leq 2$, $Q \subset B \times C : xQy \Leftrightarrow xy \geq 1$, $A = \{-1, 0, 1, 2\}$, $B = \{1, 2, 3\}$, $C = \{x \in N : x < 2\}$.

SOLUTIONS

1.
   a) $P = \{(-2, 1), (-2, 2), (-2, 3), (-1, 1), (-1, 2), (-1, 3), (0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (2, 1)\}$;

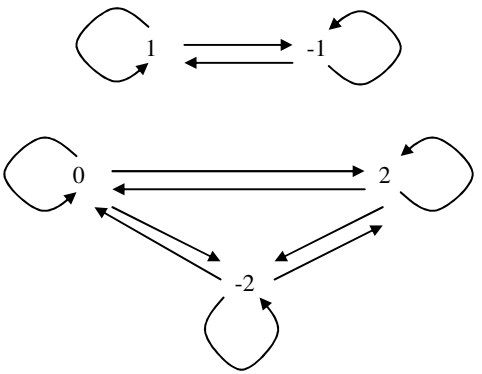
Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$P \quad \begin{matrix} 1 & 2 & 3 \end{matrix}$$
$$\begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$



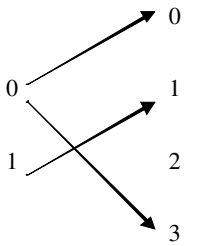
b) $P = \{(-2, -2), (-2, 0), (-2, 2), (-1, -1), (-1, 1), (0, -2), (0, 0), (0, 2), (1, -1), (1, 1), (2, -2), (2, 0), (2, 2)\}$;

$$P \quad \begin{matrix} -2 & -1 & 0 & 1 & 2 \end{matrix}$$
$$\begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$



c) $P = \{(0, 0), (0, 3), (1, 1)\}$;

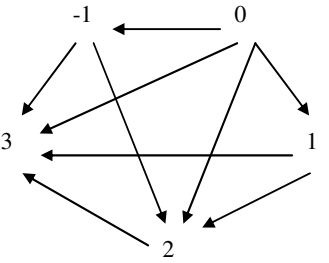
$$P \quad \begin{matrix} 0 & 1 & 2 & 3 \end{matrix}$$
$$\begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$



d) $P = \{(-1, 2), (-1, 3), (0, -1), (0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3)\}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$P \quad \begin{matrix} -1 & 0 & 1 & 2 & 3 \end{matrix}$$
$$\begin{matrix} -1 \\ 0 \\ 1 \\ 2 \\ 3 \end{matrix} \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



2. Matrices of the relations:

a)

$$P \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad Q \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad P' \quad \begin{matrix} 1 & 2 & 3 \end{matrix}$$
$$\begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Q' \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad P \cup Q \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad P \cap Q \quad \begin{matrix} 1 & 2 & 3 \end{matrix}$$
$$\begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$P \setminus Q \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad Q \setminus P \quad \begin{matrix} 1 & 2 & 3 \end{matrix} \qquad P^{-1} \quad \begin{matrix} -2 & -1 & 0 & 1 & 2 \end{matrix}$$
$$\begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{matrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$Q^{-1} \quad \begin{matrix} -2 & -1 & 0 & 1 & 2 \end{matrix}$$
$$\begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

b)

$$P=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 1 \end{array} \quad Q=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \end{array} \quad P'=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 2 & 1 & 0 & 1 & 0 \end{array}$$

$$Q'=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \end{array} \quad P\cup Q=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 & 1 \end{array}$$

$$P\cap Q=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{array} \quad P\setminus Q=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 \end{array}$$

$$Q\setminus P=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 & 0 \end{array} \quad P^{-1}=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 1 \end{array} \quad Q^{-1}=\begin{array}{c|cccc} & -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \end{array}$$

c)

$$P=\begin{array}{c|ccc} & 1 & 2 & 3 \\ -1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{array} \quad Q=\begin{array}{c|cc} & 0 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{array} \quad P'=\begin{array}{c|ccc} & 1 & 2 & 3 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 \end{array} \quad Q'=\begin{array}{c|cc} & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 0 \end{array}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$P^{-1}\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{cccc} -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array} \qquad Q^{-1}\begin{array}{c} \\ 0 \\ 1 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

3.

a) no composition exists;

b)

$$P \circ Q\begin{array}{c} \\ -1 \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cccc} -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad Q \circ P\begin{array}{c} \\ -1 \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cccc} -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

$$P^2\begin{array}{c} \\ -1 \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cccc} -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \end{array} \qquad Q^2\begin{array}{c} \\ -1 \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cccc} -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

c)

$$Q \circ P\begin{array}{c} \\ -1 \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \end{array}.$$

3.2. PROPERTIES OF THE RELATIONS

THEORY IN A NUTSHELL

In the remainder of this book you will meet a special kind of relations – **functions**.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Definition 3.3. Function
Function is such a relation that each member of the domain is related to at most one member of the counterdomain. In other words function is such a relation P for which the following condition holds: $\forall x \in D(P): \left|\{y \in D^{-1}(P): xPy\}\right| \leq 1$.

If you analyze the matrix of a relation, you can easily check whether it is function. Simply count the elements in columns – in the case of a function in each column there may be at most one 1.
A special type of function is **sequence**.

Definition 3.4. Sequence
Sequence is a function whose domain is a subset of the set of natural numbers N, i.e., such a function P, for which the following condition holds: $D(P) \subset N$.

For now, however, other properties of relations will be important. There are seven and they are presented below. In the definitions I refer to the following six matrices:

(a)
$$\begin{bmatrix} [1] & & & & \\ & [1] & & & \\ & & [1] & & \\ & & & [1] & \\ & & & & [1] \end{bmatrix}$$

(b)
$$\begin{bmatrix} [0] & & & & \\ & [0] & & & \\ & & [0] & & \\ & & & [0] & \\ & & & & [0] \end{bmatrix}$$

(c)
$$\begin{bmatrix} [\ ] & 1 & & & \\ 1 & [\ ] & & & 0 \\ & & [\ ] & & \\ & 0 & & [\ ] & \end{bmatrix}$$

(d)
$$\begin{bmatrix} [0] & 0 & & & \\ 1 & [0] & & & 0 \\ & & [0] & 1 & \\ & 0 & 0 & [0] & \end{bmatrix}$$

(e)
$$\begin{bmatrix} [\ ] & 0 & & & \\ 1 & [\ ] & & & 0 \\ & & [\ ] & 1 & \\ & 0 & 0 & [\ ] & \end{bmatrix}$$

(f)
$$\begin{bmatrix} [\ ] & 0 & & & \\ 1 & [\ ] & & & 1 \\ & & [\ ] & 1 & \\ & 1 & 0 & [\ ] & \end{bmatrix}$$

Definition 3.5. Reflexive relation
Relation $P \subset A^2$ is reflexive, if $\forall\ x \in A : xPx$.

Relation is reflexive if each element is related to itself. In the matrix of a reflexive relation all the elements of the diagonal are 1[5]. It is shown in the matrix (a).

[5] The diagonal of a matrix consists of all the elements on the line from the upper left to the bottom right corner. You will learn more about this concept in the next chapters.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Definition 3.6. Irreflexive relation
Relation $P \subset A^2$ is irreflexive, if $\forall\, x \in A : \neg\, xPx$.

Relation is irreflexive if no element is related to itself. In the matrix of an irreflexive relation all the elements of the diagonal are 0. It is shown in the matrix (b).

Definition 3.7. Symmetric relation
Relation $P \subset A^2$ is symmetric, if $\forall\, x, y \in A : xPy \Rightarrow yPx$.

Relation is symmetric if each two (not necessarily distinct) elements are related to themselves in both directions or not at all. In the matrix of a symmetric relation the elements are symmetric with respect to the diagonal: each 0 corresponds with 0, each 1 corresponds with 1. The elements on the diagonal may be arbitrary. It is shown in the matrix (c).

Definition 3.8. Antisymmetric relation
Relation $P \subset A^2$ is antisymmetric, if $\forall\, x, y \in A : xPy \Rightarrow \neg\, yPx$.

Relation is antisymmetric if each two (not necessarily distinct) elements are related to themselves in one direction or not at all. In the matrix of an antisymmetric relation the elements are not symmetric with respect to the diagonal: each 1 corresponds with 0, each 0 corresponds with 0 or 1. The elements on the diagonal must be 0. It is shown in the matrix (d).

Definition 3.9. Asymmetric relation
Relation $P \subset A^2$ is asymmetric, if $\forall\, x, y \in A : xPy \wedge yPx \Rightarrow x = y$.

Relation is asymmetric if each two distinct elements (see the difference?) are related to themselves in one direction or not at all. In the matrix of an asymmetric relation the elements are not symmetric with respect to the diagonal: each 1 corresponds with 0, each 0 corresponds with 0 or 1. The elements on the diagonal can be arbitrary. It is shown in the matrix (e).

Definition 3.10. Total relation
Relation $P \subset A^2$ is total, if $\forall\, x, y \in A : \neg\, xPy \wedge \neg\, yPx \Rightarrow x = y$.

Relation is total if each two distinct elements are related to themselves in one direction or both. In the matrix of a total relation the elements are not symmetric with respect to the diagonal: each 1 corresponds with 0 or 1, each 0 corresponds

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

with 1. The elements on the diagonal can be arbitrary. It is shown in the matrix (f).

Definition 3.11. Transitive relation
Relation $P \subset A^2$ is transitive, if $\forall\, x, y, z \in A : xPy \wedge yPz \Rightarrow xPz$.

Relation is transitive, if for each triple of elements the following holds: if the first element is related to the second one and the second one to the third one, then also the first one must be related to the third one. Relation P is transitive if and only if $P \circ P \subset P$.

EXAMPLES

Example 1
Examine the properties of the relation represented by the matrix:

$$\begin{array}{c|cccc} P & 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 0 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 4 & 0 & 0 & 0 & 1 \end{array}$$

Solution
On the diagonal there are three 0 and one 1. It means that it is neither reflexive (in this case all the elements of the diagonal should be 1) nor irreflexive (in this case all the elements of the diagonal should be 0). All the elements outside the diagonal are distributed in an asymmetric way. For example the element in the row "1" and column "2" is 1, and the element in the row "2" and column "1" is 0. The same applies to all other elements. It follows that the relation is not symmetric (in that case all the elements should be distributed symmetrically with respect to the diagonal). It is not antisymmetric, as there are some elements equal to 1 on the diagonal. The diagonal does not matter, however, in the case of asymmetry, thus this relation is asymmetric. In the matrix of a total relation, each 0 must correspond to 1 placed symmetrically with respect to the diagonal. As we can see, this is the case. Thus the relation is total. In order to verify the transitivity, we have to find the composition of P with itself:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$P^2 \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix}$$
$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Relation P is transitive if and only if $P^2 \subset P$. As we analyze the matrices, it is enough to check whether all the elements of the matrix of P^2 that are equal to 1, are present also in the matrix of P. This is the case, so P is transitive. To conclude, P is asymmetric, total and transitive. It is not reflexive, irreflexive, symmetric or antisymmetric.

Example 2
Examine the properties of the relation $P \subset A^2 : xPy \Leftrightarrow x + y \geq 5$, where $A = \{1, 2, 3, 4\}$.

Solution
The matrix of P is as follows:

$$P \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix}$$
$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

As we can see, there are 0 and 1 on the diagonal, so P is neither reflexive, nor irreflexive. All the elements are symmetrically distributed with respect to the diagonal, so it is symmetric and it is neither antisymmetric nor asymmetric. There are some symmetrically distributed zeros in the matrix, so P is not total. In order to verify the transitivity of P, we have to find the composition P^2:

$$P^2 \quad \begin{matrix} 1 & 2 & 3 & 4 \end{matrix}$$
$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

As we can see, $P^2 \not\subset P$, so P is not transitive.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
Examine the properties of the relation $P \subset A^2 : xPy \Leftrightarrow x \geq y$, where $A = \{1, 2, 3, 4\}$.

Solution
The matrix of P is as follows:

$$\begin{array}{c|cccc} P & 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

All the elements of the diagonal are 1, so P is reflexive and is not irreflexive. The elements outside the diagonal are asymmetrically distributed, so P is not symmetric and it is asymmetric (but it is not antisymmetric, as in this case all the elements of the diagonal should be 0). Each 0 corresponds to 1 on the other side of the diagonal, so P is total. The composition P^2 looks as follows:

$$\begin{array}{c|cccc} P^2 & 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 \\ 4 & 1 & 1 & 1 & 1 \end{array}$$

As $P^2 \subset P$, P is transitive.

EXERCISES

1. Examine the properties of the following relations:
   a) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow x^2 \leq y^2$;
   b) $P \subset \{-2, -1, 0, 1, 2\}^2 : xPy \Leftrightarrow x^2 \leq y^2$;
   c) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow x^3 > y^3$;
   d) $P \subset \{-2, -1, 0, 1, 2\}^2 : xPy \Leftrightarrow x + y = 0$;
   e) $P \subset \{-2, -1, 0, 1, 2\}^2 : xPy \Leftrightarrow x \cdot y \geq 0$;
   f) $P \subset \{-2, -1, 0, 1, 2\}^2 : xPy \Leftrightarrow x \cdot y < 0$;
   g) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow \exists\, k \in Z : x - y = 2k$;
   h) $P \subset \{1, 2, 3, 4, 5, 6\}^2 : xPy \Leftrightarrow \exists\, k \in Z : x^2 - y^2 = 3k$;
   i) $P \subset \{1, 2, 3, 4, 5, 6\}^2 : xPy \Leftrightarrow \exists\, k \in Z : x^2 + y^2 = 3k$;
   j) $P \subset \{1, 2, 3, 4, 5, 6\}^2 : xPy \Leftrightarrow x + y \geq 6$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. The lists contain only possessed properties:
   a) reflexive, asymmetric, total, transitive;
   b) reflexive, total, transitive;
   c) irreflexive, antisymmetric, asymmetric, total, transitive;
   d) symmetric;
   e) reflexive, symmetric, total;
   f) irreflexive, symmetric;
   g) reflexive, symmetric, transitive;
   h) reflexive, symmetric, transitive;
   i) symmetric, transitive;
   j) symmetric.

3.3. ORDERS AND PREFERENCES

THEORY IN A NUTSHELL

Ordering relations are, briefly speaking, the relations that are transitive. There are several subtypes of ordering relations. Instead of defining each type, I presented them in Fig. 3.3.

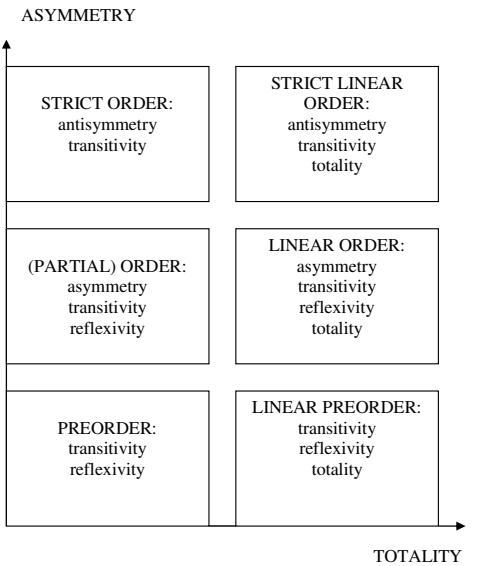


Fig. 3.3. Ordering relations

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

It is easy to see what are the dependencies between them. **Preorder** is **reflexive** and **transitive**. **Partial order** is **reflexive**, **transitive** and **asymmetric**. Finally, **strict order** is **transitive** and **antisymmetric** (i.e., also **irreflexive** and **asymmetric**). As you can see, in the case of preorders no conditions related to asymmetry apply, partial order is asymmetric and strict order – antisymmetric. Thus the conditions of asymmetry are becoming stronger when moving from one type to another. **Reflexivity** is added where it is possible (asymmetric relation cannot be reflexive, the others can). If we add **totality** to the list of properties defining preorder, partial order or strict order, we obtain, respectively, linear preorder, linear order and strict linear order.

Ordering relations are of great importance in economics, because we use them to define preferences (e.g., customers' preferences). Each of the properties occurring in the definitions of the various types of ordering relations has its interpretation:

- transitivity – if you prefer object A to object B and object B to object C, then you will certainly prefer object A to object C; this is a feature of any "reasonable" preference relation;
- antisymmetry, asymmetry – if you are comparing two different objects, you will probably prefer either one or the other of them; not all preference relations have this property however – sometimes you can find two or more objects that are equally good for you;
- totality – means that you are able to compare any two of the objects; not all the preference relations have this property; sometimes (rather rarely) it can happen that you cannot do that (note: this is not the same situation, when you are able to compare two objects and they are equally good for you);
- reflexivity – it is added to the list of the properties to make the preference relation meaningful (for example, if we define the preference relation as "xPy, when x is not worse than y", then it is obvious that each x must be related to itself, as it is not worse tha itself).

EXAMPLES

Example 1

Examine the properties of the relation represented by the following matrix and verify whether it is any kind of ordering relation.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$
\begin{array}{c|cccc}
P & 1 & 2 & 3 & 4 \\
1 & 0 & 0 & 1 & 1 \\
2 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 1 \\
4 & 0 & 0 & 0 & 0
\end{array}
$$

Solution
This relation is irreflexive, antisymmetric, asymmetric and transitive, but it is not total. It means that it is a strict order.

Example 2
The following vectors[6] are given: $a = (1, 3, 10)$, $b = (2, 7, 8)$ and $c = (5, 4, 3)$. Examine the properties of the relations $P_i \subset \{a, b, c\}^2 : xP_i y \Leftrightarrow x_i < y_i$ ($i = 1, 2, 3$). Are they ordering relations?

Solution
The notation using index i means that several relations have been defined (three in this case, as i takes the values 1, 2 and 3). Let us write each formula separately: $xP_1y \Leftrightarrow x_1 < y_1$, $xP_2 y \Leftrightarrow x_2 < y_2$, $xP_3 y \Leftrightarrow x_3 < y_3$. Two vectors are P_1-related, when their first coordinates satisfy the respective inequality (thus, in order to check whether the relation occurs, we compare the numbers 1, 2 and 5). Similarly in the case of P_2 we are interested in second coordinates (i.e., 3, 7 and 4), and in the case of P_3 – third ones (i.e., 10, 8 i 3). The matrices of the relations look as follows:

$$
\begin{array}{c|ccc}
P_1 & a & b & c \\
a & 0 & 1 & 1 \\
b & 0 & 0 & 1 \\
c & 0 & 0 & 0
\end{array}
\qquad
\begin{array}{c|ccc}
P_2 & a & b & c \\
a & 0 & 1 & 1 \\
b & 0 & 0 & 0 \\
c & 0 & 1 & 0
\end{array}
\qquad
\begin{array}{c|ccc}
P_3 & a & b & c \\
a & 0 & 0 & 0 \\
b & 1 & 0 & 0 \\
c & 1 & 1 & 0
\end{array}
$$

As you can easily check, all three relations are irreflexive, antisymmetric, asymmetric, transitive and total, so each of them is a strict linear order.

[6] You will learn more about vectors soon. Now you must only know that each vector may be interpreted as a finite sequence of numbers, and its elements are denoted by the same letter as the vector itself, with respective index. In this example one of the vectors, a, has three elements (we call them also coordinates), namely $a_1 = 1$, $a_2 = 3$ and $a_3 = 10$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
In the Table 3.1 are given: the values of expected return (million PLN) and risk[7] (%) for four possible investments. Define appropriate preference relations in the given set of investments (present their formulae and matrices).

Table 3.1. Information on the investments

Investment	A	B	C	D
Return	4	4	2	5
Risk	25	15	15	20

Solution
Let E denote the expected return on investment. We want it to be the maximum. In other words, the investment X is better than the investment Y if the expected return on X is greater than the expected return on Y. To address the situation, when the return of two investments is the same, we can formulate it this way: investment X is not worse than the investment Y when the expected return on X is not less than the expected return on Y. If this relation is denoted by P_E, we can define it with the following formula:

$$P_E \subset \{A, B, C, D\}^2 : XP_EY \Leftrightarrow E(X) \geq E(Y).$$

Its matrix looks as follows:

$$\begin{array}{c} P_E \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

On the other hand, the lower is the risk, the better is the investment. If we denote the risk by S and the respective relation by P_S, then we can define it with the formula:

$$P_S \subset \{A, B, C, D\}^2 : XP_SY \Leftrightarrow S(X) \leq S(Y).$$

[7] You will find more about the expectation and risk in Chapter 14 (Probability). Now it is only important that a reasonable way of selection is to maximize the expected return and to minimize the risk.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Its matrix looks as follows:

$$\begin{array}{c} P_S \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

What remained is to verify whether those two relations are indeed preference relations, i.e., whether they are ordering relations. The answer is: yes. They are both reflexive, transitive and total, so they are both linear preorders.

EXERCISES

1. Examine the properties of the following relations and decide whether they are any kind of ordering relation:
   a) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow x^2 \geq y^2$;
   b) $P \subset \{-2, -1, 0, 1, 2\}^2 : xPy \Leftrightarrow x^2 \geq y^2$;
   c) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow x^3 < y^3$;
   d) $P \subset \{1, 2, 3, 4\}^2 : xPy \Leftrightarrow \exists\, k \in Z : x - y = 3k$.
2. Consider the set of vectors $A = \{a, b, c, d\}$, where $a = (1, 2, 7)$, $b = (2, 4, 9)$, $c = (1, 3, 10)$, $d = (2, 3, 8)$. Examine the properties of the following relations and decide whether they are any kind of ordering relation.
   a) $xP_i y \Leftrightarrow x_i > y_i$ $(i = 1, 2, 3)$;
   b) $xP_i y \Leftrightarrow x_i \leq y_i$ $(i = 1, 2, 3)$;
   c) $xP_i y \Leftrightarrow x_i = y_i$ $(i = 1, 2, 3)$.
3. Investments A, B, C and D are described with pairs of numbers: each of them (denoted by X) is represented by the expected return $E(X)$ and risk $S(X)$, written in the form of vector $X = (E(X), S(X))$. Present the matrices of the preference relations defined as $XP_E Y \Leftrightarrow E(X) \geq E(Y)$ and $XP_S Y \Leftrightarrow S(X) \leq S(Y)$.
   a) $A = (20, 10)$, $B = (30, 20)$, $C = (35, 20)$, $D = (40, 25)$;
   b) $A = (30, 10)$, $B = (20, 5)$, $C = (25, 10)$, $D = (20, 5)$;
   c) $A = (5, 10)$, $B = (7, 15)$, $C = (3, 6)$, $D = (4, 5)$.

SOLUTIONS

1. The lists contain only possessed properties:
   a) reflexive, asymmetric, total, transitive – linear order;
   b) reflexive, total, transitive – linear preorder;
   c) irreflexive, antisymmetric, asymmetric, total, transitive – strict linear order;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

d) reflexive, symmetric, transitive – isn't any type of ordering relation.

2. The lists contain only possessed properties:
   a) all three relations irreflexive, antisymmetric, asymmetric and transitive; moreover P_3 is total; P_1 and P_2 – strict orders; P_3 – strict linear orders;
   b) all three relations reflexive, total and transitive; moreover P_3 is asymmetric; P_1 and P_2 – linear preorders; P_3 – linear order;
   c) all three relations are reflexive, symmetric and transitive; moreover P_3 is asymmetric; none of them is an ordering relation.
3. Matrices of the relations:
   a)

$$P_E = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad P_S = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array};$$

   b)

$$P_E = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \end{array} \qquad P_S = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array};$$

   c)

$$P_E = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array} \qquad P_S = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}.$$

3.4. PARETO EFFICIENCY

THEORY IN A NUTSHELL

In the real life it happens very often that the preferences are not defined in a simple, clear way.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Imagine for instance that you want to buy a car. On the one hand matters to you its size (you want to buy as large car as possible), on the other hand, price (the cheaper, the better), moreover, fuel consumption, engine power, color, … Exactly. There is rather no car that would be the best from each point of view.
Another situation: you are going with friends to the movies. But each of you has different preferences for movies. And what will you do?
Described situations are **multicriteria** (or **multiobjective**) problems. We call them so, since the decision has to be taken on the basis of at least two different criteria. Usually it will be impossible to choose one solution, the best in every respect. However, you can reduce their number by focusing on the **Pareto set**, also called the set of **efficient solutions**. Each element of the Pareto set is called the **Pareto optimum** or **Pareto-optimal (Pareto-efficient) solution (decision)**.

Definition 3.12. Dominated solution
Suppose that a multicriteria problem is given. The solution Y is dominated by the solution X, if X is not worse than Y in all respects, and is better with respect to at least one criterion.

Definition 3.13. Pareto set
Pareto set is the set of solutions that are not dominated by any other solution.

Let us recall the "car" example. Assume that the only two criteria that matter are the fuel consumption and the engine power. If two cars have the same power of 150 hp[8], but one of them consumes 7 l of fuel per 100 km, and another one 8 l per 100 km, it means that the second one is dominated (the first car is better with respect to consumption and not worse with respect to power). For that reason, the second car is not a member of the Pareto set.
You do not have to derive the Pareto set directly from the definition. You can use the following procedure:

1. Find the intersection of all the considered preference relations. Let us denote it by M.
2. Find the relation $S = (M^{-1})' \cap M$. It is the strict order associated with the preorder M (all the initial preference relations can be any kind of ordering relations, and so can be M). If you use the matrices, in order to find S it is enough to change all the symmetrically distributed 1 into 0, as well as fill the diagonal with 0.
3. Find all the **maximal elements**, i.e. all the elements which no other element is related to. In the matrix it is enough to find all the elements whose columns contain only 0. These elements are members of the Pareto set.

[8] A unit of power, *horsepower*. 1 hp = 746 W.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1

The following vectors are given: $a = (1, 3, 5)$, $b = (2, 3, 8)$ and $c = (5, 4, 8)$, as well as the following relations: $P_i \subset \{a, b, c\}^2 : xP_i y \Leftrightarrow x_i \leq y_i$ $(i = 1, 2, 3)$. Find the Pareto set.

Solution

The matrices of the relations are as follows:

$$P_1 = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \qquad P_2 = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \qquad P_3 = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

In order to find the Pareto set, we begin with finding the intersection of all considered relations. Recall that this means saving only those elements equal to 1, which are present in all the matrices. In this case it looks as follows (remember that we denote it by M):

$$M = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The next step is finding the strict order, $S = (M^{-1})' \cap M$ associated with M. Recall that in order to do this, we remove all the elements equal to 1 from the diagonal and all such elements distributed symmetrically in the matrix (there is no such pair in this example). The result looks as follows:

$$S = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

The last step is to find the maximal elements, i.e. the elements corresponding with the columns, where all the elements are 0. There is only one such element: a. Thus the Pareto set is $P = \{a\}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 2
Table 3.2 presents the values of the expected return (million PLN) and the associated risks (%) for four investments. Determine the set of effective decisions.

Table 3.2. Information on the investments

Investment	A	B	C	D
Return	5	5	3	4
Risk	15	15	10	20

Solution
Let us define the appropriate relations. They are defined with the formulae $XP_EY \Leftrightarrow E(X) \geq E(Y)$ and $XP_SY \Leftrightarrow S(X) \leq S(Y)$, where $P_E, P_S \subset \{A, B, C, D\}^2$. Their matrices look as follows:

$$P_E = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array} \qquad P_S = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

The intersection and respective strict order are presented below (observe that this time there are symmetric 1: they represent the pairs (A, B) and (B, A); it was thus necessary to remove them both).

$$M = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \qquad S = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

As we can see, there are elements equal to 1 in the column "D", so the Pareto set consists of three remaining investments: $P = \{A, B, C\}$.

EXERCISES

1. Given the set of vectors $A = \{a, b, c, d\}$, where $a = (1, 2, 7)$, $b = (2, 4, 9)$, $c = (1, 3, 10)$, $d = (2, 3, 8)$, find the Pareto set:
   a) $xP_iy \Leftrightarrow x_i > y_i$ $(i = 1, 2, 3)$;
   b) $xP_iy \Leftrightarrow x_i \leq y_i$ $(i = 1, 2, 3)$;
   c) $xP_iy \Leftrightarrow x_i < y_i$ $(i = 1, 2, 3)$;
   d) $xP_iy \Leftrightarrow x_i \geq y_i$ $(i = 1, 2, 3)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Investments A, B, C and D are described with pairs of numbers: each of them (denoted by X) is represented by the expected return $E(X)$ and risk $S(X)$, written in the form of vector $X = (E(X), S(X))$. The preference relations are defined as $XP_EY \Leftrightarrow E(X) \geq E(Y)$ and $XP_SY \Leftrightarrow S(X) \leq S(Y)$. Find the set of efficient decisions:
   a) $A = (20, 10)$, $B = (30, 20)$, $C = (35, 20)$, $D = (40, 25)$;
   b) $A = (30, 10)$, $B = (20, 5)$, $C = (25, 10)$, $D = (20, 5)$;
   c) $A = (5, 10)$, $B = (7, 15)$, $C = (3, 6)$, $D = (4, 5)$.
3. Employees A, B, C and D have been assessed due to three criteria: the number of projects carried out in the past year L (the higher, the better), the number of absences in the past year N (the lower, the better) and the monthly payment in thousand PLN (should be as low as possible). Define the appropriate relations, determine the Pareto set and decide whether using only this method the owner can decide who should be fired, if they want to run a fifty percent reduction in employment. The data are given in the form of vectors $X = (L(X), N(X), W(X))$.
   a) $A = (10, 5, 3)$, $B = (9, 7, 4)$, $C = (9, 3, 4)$, $D = (8, 3, 3)$;
   b) $A = (7, 3, 4)$, $B = (7, 7, 4)$, $C = (4, 3, 5)$, $D = (4, 3, 3)$;
   c) $A = (10, 7, 5)$, $B = (9, 6, 5)$, $C = (12, 5, 4)$, $D = (10, 4, 4)$.

SOLUTIONS

1. The Pareto sets:
   a) $P = \{b, c, d\}$;
   b) $P = \{a\}$;
   c) $P = \{a, c\}$;
   d) $P = \{b, c\}$.
2. The sets of efficient decisions:
   a) $P = \{A, C, D\}$;
   b) $P = \{A, B, D\}$;
   c) $P = \{A, B, D\}$.
3. In all three cases, the formulae of the ordering relations are as follows: $XP_LY \Leftrightarrow L(X) \geq L(Y)$, $XP_NY \Leftrightarrow N(X) \leq N(Y)$ and $XP_NY \Leftrightarrow N(X) \leq N(Y)$, where $P_L, P_N, P_W \subset \{A, B, C, D\}^2$. The remainders of the solutions:
   a) $P = \{A, C, D\}$, it is impossible to indicate the two worst employees, it is necessary to use another method;
   b) $P = \{A, D\}$, so the employees B and C are clearly worse and should be fired;
   c) $P = \{C, D\}$, so the employees A and B are clearly worse and should be fired.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 4
MATRICES AND VECTORS

4.1. BASIC CONCEPTS. MATRIX ALGEBRA

THEORY IN A NUTSHELL

Matrix is a rectangular table filled with numbers. We already considered a special kind of matrices in the previous chapter. Now you are going to learn a little bit more.
Matrices are denoted by capital letters A, B, … A special type of matrix with only one row or only one column is called a vector (row vector or column vector, respectively). The vectors are denoted by lowercase letters a, b, … Matrices are bounded with square brackets, vectors can written in the square, but also round brackets. Examples of matrices (including vectors) can be found below:

$$A=\begin{bmatrix}1 & 2 & 7\\ 5 & 1 & -4\end{bmatrix},\quad B=\begin{bmatrix}-1 & 2\\ -8 & 3\\ 3.5 & 0\end{bmatrix},\quad a=\begin{pmatrix}1\\ -2\\ 7\end{pmatrix},\quad I_3=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},\quad C=\begin{bmatrix}3 & 0\\ 0 & 1\\ 3.5 & 0\end{bmatrix}.$$

We say that the matrix is **square** when it has the same number of rows as columns. A special type of matrix is the **identity matrix**. It is a matrix whose elements on the main diagonal are all ones, all the remaining zeros. We denote it with the letter I, and if we know its dimensions – add the index indicating the number of rows (and columns at the same time). The second last matrix among the matrices presented above is the identity matrix of three rows, hence we denote it by I_3.
Each element of the matrix has its **coordinates**; first of them is the row index and the second – the column index. The coordinates are obviously positive integers. Any chosen element of the matrix can be written by using the symbol of the matrix and the brackets, for example in the case of matrices A and a we have: $A[1, 1] = 1$, $A[1, 2] = 2$, $A[1, 3] = 7$, $A[2, 1] = 5$, $A[2, 2] = 1$, $A[2, 3] = –4$, $a[1, 1] = 1$, $a[2, 1] = –2$ and $a[3, 1] = 7$. As a is a column vector, we can skip the index of the column: $a[1] = 1$, $a[2] = –2$ and $a[3] = 7$ (the same can be done with a row vector). If the matrix is not named, then we skip the letter and refer to the elements by giving the right indices in the brackets (i.e., element [1, 2], element [3] and so on). There is also another notation – much more popular. In the case of a matrix denoted by an uppercase letter, in order to indicate a specific element we use respective lowercase letter and add the lower indices (first row, then column). For example, in the case of matrix A we have: $a_{11} = 1$, $a_{12} = 2$, $a_{13} = 7$,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$a_{21} = 5$, $a_{22} = 1$, $a_{23} = -4$. When we consider a vector, then we use the same letter, with only one lower index, which would result with $a_1 = 1$, $a_2 = -2$ and $a_3 = 7$ in the case of vector a. We will use both notations, depending on the context.

We can perform some of the standard arithmetic operations on the matrices:

- in order to **add or subtract** two matrices, we add or subtract the elements having same coordinates; both matrices must have same numbers of rows and columns;
- to **multiply a matrix by a number**, we multiply each element of the matrix by that number;
- to **multiply one matrix by another**, we multiply each row of the first matrix by each column of the other one; multiplying the row and the column means that we multiply one by one their elements (the first with the first, the second with the second and so on), and sum up the products obtained this way; to make multiplication practicable, the first matrix must have same number of columns as the number of rows of the second one;
- in order to **transpose a matrix** (which we denote by A^T), we copy the elements of rows as the elements of columns of the resulting matrix (or equivalently we can copy the elements of columns as the elements of rows)[9].

In order to implement the mentioned operations in **WolframAlpha®**, you need to use the following notation. The matrices are represented by the sequences of rows, placed between curly brackets "{" and "}" and separated with commas (","), where each row is again a sequence formed in the same way, with the elements being the numbers. The addition and subtraction are denoted with "+" and "–", respectively, the matrix multiplication with a dot (".") and the multiplication of a matrix and a number with ordinary multiplication symbol (asterisk, "*"). In order to transpose a matrix, it is necessary to use the command "transpose". For instance, to transpose the matrix A defined above, you should type:

"transpose {{1,2,7},{5,1,-4}}",

while the multiplication of A and C should be implemented by typing:

"{{1,2,7},{5,1,-4}}.{{3,0},{0,1},{3.5,0}}".

[9] If you associate it with inversing the relation, it's very good, because by transposing the matrix of a relation, we determine the matrix of the inverse relation. Be careful with that, because soon you will learn the concept of the inverse matrix, which has nothing in common with the inverse relation!

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1
Perform the following operations on the matrices from the previous page: $2A$, B^T, $B - C$, $B + C$.

Solution
To multiply a matrix by a number, we multiply all the elements of the matrix by this number. Thus, the elements of $2A$ are equal to: $2 \cdot 1 = 2$, $2 \cdot 2 = 4$, $2 \cdot 7 = 14$, $2 \cdot 5 = 10$, $2 \cdot 1 = 2$, $2 \cdot (-4) = -8$. Finally, the result is as follows:

$$2A = \begin{bmatrix} 2 & 4 & 14 \\ 10 & 2 & -8 \end{bmatrix}.$$

Transposition of a matrix is to change the rows into columns (or vice versa), what means that B^T looks as follows:

$$B^T = \begin{bmatrix} -1 & -8 & 3.5 \\ 2 & 3 & 0 \end{bmatrix}.$$

To determine the sum or difference of the matrix, we determine the sum or difference of their individual components. Thus, elements of difference $B - C$ are: $-1 - 3 = -4$, $2 - 0 = 2$, $-8 - 0 = -8$, $3 - 1 = 2$, $3.5 - 3.5 = 0$, $0 - 0 = 0$, while the elements of the sum $B + C$ are: $-1 + 3 = 2$, $2 + 0 = 2$, $-8 + 0 = -8$, $3 + 1 = 4$, $3.5 + 3.5 = 7$, $0 + 0 = 0$. Both resulting matrices are presented below:

$$B - C = \begin{bmatrix} -4 & 2 \\ -8 & 2 \\ 0 & 0 \end{bmatrix}, \quad B + C = \begin{bmatrix} 2 & 2 \\ -8 & 4 \\ 7 & 0 \end{bmatrix}.$$

Example 2
Perform the following operations on the matrices from the page 68: Aa, AB, BA.

Solution
Let us start with the product Aa. To determine the shape of the resulting matrix, then the elements, we can use the **Falk Scheme**. To do this, write the first matrix to the left of the resulting matrix, and the second – above (Fig. 4.1a). The resulting matrix has as many rows as the first matrix multiplied, and as many

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

columns as the second. In our case, this means two rows and one column (all items marked with asterisks). To determine the value of any element of the resulting matrix (say, first), we multiply the line on its left (i.e. the row having elements 1, 2 and 7) by the column placed above (i.e. the column having elements 1, –2 and 7). This way we obtain: $1\cdot 1 + 2\cdot(-2) + 7\cdot 7 = 46$ (Fig. 4.1b). Similarly, in order to find the second element, we multiply the line on its left (i.e. the row having elements: 5, 1 and –4) by the column placed above (i.e. the column having elements: 1, –2 and 7) and we get: $5\cdot 1 + 1\cdot(-2) + (-4)\cdot 7 = -25$ (Fig. 4.1c).

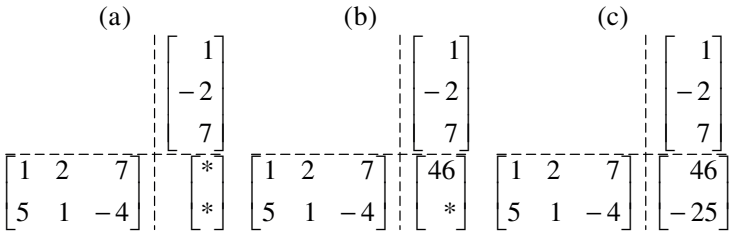


Fig. 4.1. Matrix multiplication using the Falk Scheme

Finally, the resulting matrix has the form:

$$Aa = \begin{bmatrix} 46 \\ -25 \end{bmatrix}.$$

To perform the operation AB, we proceed in exactly the same way. The difference is that this time the resulting matrix will have two columns. The corresponding diagram (only output and the final figure) is shown in Fig. 4.2. To determine the elements of the resulting matrix, it is necessary to perform actions on the corresponding rows and columns:

$1\cdot(-1) + 2\cdot(-8) + 7\cdot 3.5 = 7.5,$

$1\cdot 2 + 2\cdot 3 + 7\cdot 0 = 8,$

$5\cdot(-1) + 1\cdot(-8) + (-4)\cdot 3.5 = -27$

and

$5\cdot 2 + 1\cdot 3 + (-4)\cdot 0 = 13.$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

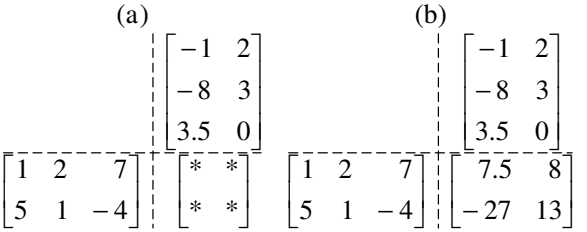


Fig. 4.2. Matrix multiplication using the Falk Scheme

The product of A and B is thus equal to:

$$AB = \begin{bmatrix} 7.5 & 8 \\ -27 & 13 \end{bmatrix}.$$

Following a procedure analogous to two previous cases, you can also get the value of the product of the matrices B and A (Falk Scheme and final result in Fig. 4.3).

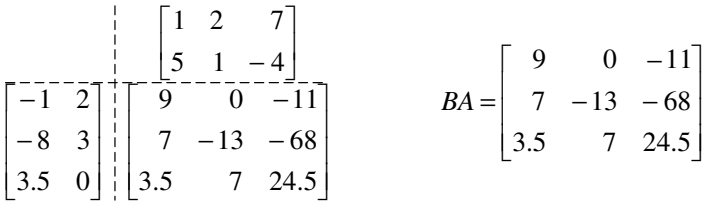


Fig. 4.3. Matrix multiplication using the Falk Scheme

Observe that **the matrix multiplication is not commutative**: in the current example, the matrices AB and BA have even different sizes (2×2 and 3×3, respectively). That's why it is so important not to change the order of multiplication in the formulae.

EXERCISES

1. Given the following matrices: $A = \begin{bmatrix} 1 & 7 & 3 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -2 & 1 & -1 \end{bmatrix}$, $C = \begin{bmatrix} 2 & -2 & 5 & 0 \\ 1 & 0 & 1 & 5 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 6 & -5 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$, perform the following operations:
   a) A^T, B^T, C^T, D^T;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

b) $2A$, $3B$, $-4C$, $-5D$;
c) $A + B$, $A - B$, $C + D$, $C - D$;
d) $A^T \cdot B$, $B \cdot A^T$, $C \cdot A^T$, $B \cdot D^T$.

SOLUTIONS

1. a) $A^T = \begin{bmatrix} 1 \\ 7 \\ 3 \\ -2 \end{bmatrix}$, $B^T = \begin{bmatrix} 3 \\ -2 \\ 1 \\ -1 \end{bmatrix}$, $C^T = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 5 & 1 \\ 0 & 5 \end{bmatrix}$, $D^T = \begin{bmatrix} 5 & -2 \\ 6 & 0 \\ -5 & 0 \\ 0 & 2 \end{bmatrix}$;

b) $2A = \begin{bmatrix} 2 & 14 & 6 & -4 \end{bmatrix}$, $3B = \begin{bmatrix} 9 & -6 & 3 & -3 \end{bmatrix}$,
$-4C = \begin{bmatrix} -8 & 8 & -20 & 0 \\ -4 & 0 & -4 & -20 \end{bmatrix}$,
$-5D = \begin{bmatrix} -25 & -30 & 25 & 0 \\ 10 & 0 & 0 & -10 \end{bmatrix}$;

c) $A + B = \begin{bmatrix} 4 & 5 & 4 & -3 \end{bmatrix}$, $A - B = \begin{bmatrix} -2 & 9 & 2 & -1 \end{bmatrix}$,
$C + D = \begin{bmatrix} 7 & 4 & 0 & 0 \\ -1 & 0 & 1 & 7 \end{bmatrix}$, $C - D = \begin{bmatrix} -3 & -8 & 10 & 0 \\ 3 & 0 & 1 & 3 \end{bmatrix}$;

d) $A^T B = \begin{bmatrix} 3 & -2 & 1 & -1 \\ 21 & -14 & 7 & -7 \\ 9 & -6 & 3 & -3 \\ -6 & 4 & -2 & 2 \end{bmatrix}$, $BA^T = \begin{bmatrix} -6 \end{bmatrix}$,
$CA^T = \begin{bmatrix} 3 \\ -6 \end{bmatrix}$, $BD^T = \begin{bmatrix} -2 & -8 \end{bmatrix}$.

4.2. ELEMENTARY OPERATIONS

THEORY IN A NUTSHELL

In the remainder of the book you will learn how to apply the matrices to solve various kinds of problems. In addition to the activities they met in the previous section you also need a more complex tool, which is the **Gauss–Jordan elimination (GJ elimination)**, consisting of elementary operations. Your goal will be each time to transform the matrix in such a way that in the resulting matrix there are as many columns of the identity matrix of respective size as

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

possible (this final form is called the **reduced row echelon form**). You will use the following elementary operations:

- swapping the position of two rows; the notation $w_i \leftrightarrow w_j$ means that you swap rows *i* and *j*;
- multiplying row by a nonzero scalar; the notation $w_j' = c \cdot w_j$ means that you multiply each element of the row w_j by the number *c*;
- subtract from one row a scalar multiple of another; the notation $w_j' = w_j - c \cdot w_i$ means that each element of the row w_j dis decreased by the product of the respective element of the row w_i and the number *c*.

Sometimes (in chosen applications) it is allowed to perform same operations on the columns. You will not need, however, to act this way.
In order to obtain different columns of the identity matrix, the row operations must be performed in such a way that in each column of the initial matrix, one of the elements becomes 1 and other ones 0. You will learn how to do it by studying the examples.
In order to implement the Gauss–Jordan elimination in **WolframAlpha®**, you have to use the command "row reduce". For example, if you want to reduce the matrix from the previous section

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 5 & 1 & -4 \end{bmatrix},$$

then you need the following code:

"row reduce {{1,2,7},{5,1,-4}}".

EXAMPLES

Example 1
Assume that the following matrix is given:

$$\begin{bmatrix} 2 & 5 & 2 \\ 3 & 0 & 1 \end{bmatrix}.$$

Perform the following elementary operations:

$$w_1 \leftrightarrow w_2,\ w_1' = \frac{1}{3} w_1,\ w_2' = w_2 - 2w_1.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

Let us start with the first operation: $w_1 \leftrightarrow w_2$. This notation means that we are supposed to swap the first row with the second one. The result of this operation is presented in the matrix (a):

$$\underset{\text{(a)}}{\begin{bmatrix} 3 & 0 & 1 \\ 2 & 5 & 2 \end{bmatrix}} \quad \underset{\text{(b)}}{\begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 2 & 5 & 2 \end{bmatrix}} \quad \underset{\text{(c)}}{\begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 5 & \frac{4}{3} \end{bmatrix}}$$

The second operation is the multiplication of the first row by 1/3: $w_1' = 1/3 w_1$. In order to do so, we multiply each element of this row by 1/3: $3 \cdot 1/3 = 1$, $0 \cdot 1/3 = 0$, $1 \cdot 1/3 = 1/3$. The result is presented in matrix (b).

The last operation is the subtraction of twice the first row from the second one: $w_2' = w_2 - 2w_1$. In order to do that, we subtract from each element of the second row, the respective (staying in the same column) element of the first row, multiplied by 2: $2 - 2 \cdot 1 = 0$, $5 - 2 \cdot 0 = 5$, $2 - 2 \cdot 1/3 = 4/3$. The result of this last operation is presented in matrix (c).

Example 2

The following matrix is given:

$$\begin{bmatrix} 2 & 3 & 4 & 7 \\ -3 & 3 & 5 & 1 \\ 5 & 7 & 9 & 12 \end{bmatrix}$$

Perform simultaneously the following elementary operations:

$$w_1' = \frac{1}{2} w_1, \; w_2' = w_2 + \frac{3}{2} w_1, \; w_3' = w_3 - \frac{5}{2} w_1.$$

Solution

This time the elementary operations have to be performed simultaneously and therefore more expressions will refer to the input matrix, and not to the matrices being the intermediate calculation results. This way of performing elementary operations accelerates the elimination process. The only operation that should not be done simultaneously with others, is swapping the position of two rows.

Let us start with multiplying the first row by 1/2 ($w_1' = 1/2 w_1$): $1/2 \cdot 2 = 1$, $1/2 \cdot 3 = 3/2$, $1/2 \cdot 4 = 2$, $1/2 \cdot 7 = 7/2$. The result of this operation can be found

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

in matrix (a) below. Second operation is adding the first row multiplied by 3/2 to the second row ($w_2' = w_2 + 3/2 w_1$): $-3 + 3/2 \cdot 2 = 0$, $3 + 3/2 \cdot 3 = 15/2$, $5 + 3/2 \cdot 4 = 11$, $1 + 3/2 \cdot 7 = 23/2$. The result of the two first operations can be found in matrix (b). The last operation is subtracting the first row multiplied by 5/2 from the third row ($w_3' = w_3 - 5/2 w_1$): $5 - 5/2 \cdot 2 = 0$, $7 - 5/2 \cdot 3 = -1/2$, $9 - 5/2 \cdot 4 = -1$, $12 - 5/2 \cdot 7 = -11/2$. You can find the final result in matrix (c).

(a)
$$\begin{bmatrix} 1 & \frac{3}{2} & 2 & \frac{7}{2} \\ & & & \\ & & & \end{bmatrix}$$

(b)
$$\begin{bmatrix} 1 & \frac{3}{2} & 2 & \frac{7}{2} \\ 0 & \frac{15}{2} & 11 & \frac{23}{2} \\ & & & \end{bmatrix}$$

(c)
$$\begin{bmatrix} 1 & \frac{3}{2} & 2 & \frac{7}{2} \\ 0 & \frac{15}{2} & 11 & \frac{23}{2} \\ 0 & -\frac{1}{2} & -1 & -\frac{11}{2} \end{bmatrix}$$

Observe that after performing the three preset operations, the first column became a column of the identity matrix. A coincidence? Not at all. It was intentional. But how to select the operations and their coefficients in order to generate such a column?

Recall that each column of the identity matrix consists of one 1 and all its remaining elements are 0. In order to obtain a 1, it is necessary to divide the respective row by the element which is supposed to become 1 (or, in other words to multiply by its reciprocal). As we wanted to obtain 1 at the position [1, 1] (i.e., change 2 to 1), it was necessary to divide the whole row by 2 (thus the operation: $w_1' = 1/2 w_1$). The element that turns into 1 is called the **pivot element** (its row and column are called the **pivot row** and the **pivot column**, respectively). All the other elements of the pivot column should become 0. For this purpose, we subtract from each of those rows a multiplicity of the pivot row. The coefficient always equals to the zeroed element divided by the pivot element. Thus turning the element [2, 1] (equal to –3) into 0 needs subtracting the first row multiplied by –3/2 from the second row (–3 – zeroed element, 2 – pivot element). And as subtracting –3/2 is exactly the same as adding 3/2, the elementary operation had the form $w_2' = w_2 + 3/2 w_1$. Similarly, in order to zero the element [3, 1] (equal to 5), it was necessary to subtract the first row multiplied by 5/2 from the second row (zeroed element 5, pivot element 2). The operation had the form $w_3' = w_3 - 5/2 w_1$.

In order to reduce the matrix and obtain its reduced row echelon form, it is necessary to repeat such a sequence of elementary operations for all rows and columns, i.e. reduce the first column using the first row (what we already have done), then reduce the second column using the second row, then reduce the

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

third column using the third row and so on. If at some step it is impossible (the pivot element is 0 and the pivot row cannot be divided by it), then we swap the pivot row with some row placed below. If there is no row below with a nonzero element in the pivot column, we skip this column and move forward to another one. Remember to move the rows consisting only of zeros to the bottom of the matrix each time when such a row appears. We finish at the moment when it is no longer possible to reduce any column.

Example 3
Reduce the following matrix to the reduced row echelon form:

$$\begin{bmatrix} 2 & 4 & 4 & 8 \\ 3 & 6 & 7 & 8 \\ 4 & 10 & 5 & 7 \end{bmatrix}.$$

Solution
In the first step, the first column of the matrix is the pivot column and the first row is the pivot row. The element [1, 1] equal to 2 is thus the pivot element. It has to be changed to 1, thus the operation on the pivot row takes the form $w_1' = 1/2 w_1$. The remaining elements of the pivot column must become 0, so we subtract the respective multiplicities of the pivot (first) row from the remaining rows. Let us recall that these multiplicities are equal to the ratios of the zeroed elements (the elements of the pivot column) and the pivot element, thus 3/2 and 4/2). The respective operations are: $w_2' = w_2 - 3/2 w_1$ and $w_3' = w_3 - 2w_1$. The result is as follows (the pivot element is marked with square brackets):

$$\begin{bmatrix} [2] & 4 & 4 & 8 \\ 3 & 6 & 7 & 8 \\ 4 & 10 & 5 & 7 \end{bmatrix} \begin{matrix} w_1' = \frac{1}{2} w_1 \\ w_2' = w_2 - \frac{3}{2} w_1 \\ w_3' = w_3 - 2w_1 \end{matrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 0 & 1 & -4 \\ 0 & 2 & -3 & -9 \end{bmatrix}.$$

Note how the operations are written (each operation next to respective row). Moreover, the arrow indicates the order of the matrices. It is not a formal rule, but it makes keeping everything in order much easier.
In the next step, the second column is the pivot column. The pivot element is [2, 2], equal to 0. As it cannot be changed into 1 (the division by 0 is forbidden), it is necessary to swap the pivot row with any row lying below. In this case we can use only the third row (the pivot element is marked with square brackets):

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & [0] & 1 & -4 \\ 0 & 2 & -3 & -9 \end{bmatrix} \quad w_2 \leftrightarrow w_3 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 2 & -3 & -9 \\ 0 & 0 & 1 & -4 \end{bmatrix}.$$

Now it is possible to divide and subtract the rows (the pivot element equals to 2). To be more specific, we divide the second row by 2, subtract the second row from the first one (i.e., subtract the second row multiplied by 2/2), and the third row remains unchanged (i.e., we subtract the second row multiplied by 0/2). The result is as follows (the pivot element is marked with square brackets):

$$\begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & [2] & -3 & -9 \\ 0 & 0 & 1 & -4 \end{bmatrix} \begin{matrix} w_1' = w_1 - w_2 \\ w_2' = \frac{1}{2} w_2 \\ w_3' = w_3 \end{matrix} \rightarrow \begin{bmatrix} 1 & 0 & 5 & 13 \\ 0 & 1 & -\frac{3}{2} & -\frac{9}{2} \\ 0 & 0 & 1 & -4 \end{bmatrix}.$$

In the last step we eliminate the third column using the third row. As the pivot element [3, 3] equals to 1 and the remaining elements of the pivot column are 5 and –3/2, the operations will be as follows (the pivot element is marked with square brackets).

$$\begin{bmatrix} 1 & 0 & 5 & 13 \\ 0 & 1 & -\frac{3}{2} & -\frac{9}{2} \\ 0 & 0 & [1] & -4 \end{bmatrix} \begin{matrix} w_1' = w_1 - 5w_3 \\ w_2' = w_2 + \frac{3}{2} w_3 \\ w_3' = w_3 \end{matrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 33 \\ 0 & 1 & 0 & -\frac{21}{2} \\ 0 & 0 & 1 & -4 \end{bmatrix}.$$

Example 4
Reduce the following matrix to the reduced row echelon form:

$$\begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 11 \end{bmatrix}.$$

Solution
We begin with eliminating the first column. For this purpose we have to perform the following operations (the pivot element is marked with square brackets):

$$\begin{bmatrix} [1] & 2 & 5 \\ 2 & 4 & 11 \end{bmatrix} \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \end{matrix} \rightarrow \begin{bmatrix} 1 & 2 & 5 \\ 0 & (0) & 1 \end{bmatrix}.$$

Now the pivot element should be [2, 2] (marked with round brackets). It is however equal to 0. For that reason we would like to swap the second row with

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

another one, lying below. But there is no such row. In such a case we skip the column and consider next one to the right as the pivot column (so the pivot element moves respectively). We proceed in the same way if there are some rows below the pivot row, but all the elements below the pivot element are 0:

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 0 & \boxed{1} \end{bmatrix} \begin{matrix} w_1' = w_1 - 5w_2 \\ w_2' = w_2 \end{matrix} \quad \rightarrow \quad \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \; 2 \; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right].$$

The pivot element has been marked with square brackets, and the identity submatrix in the resulting reduced matrix – with round brackets. Note that its columns are not consecutive – it is not necessary.

Note. Sometimes you may want to swap the rows to make the calculations easier (for example, there is a 1 somewhere in the pivot column below the pivot element, so after swapping the rows you will not have to operate with fractions). In this book, we will not perform this type of operations, so if you decide to do that, you may find that certain intermediate matrices are different than those in the book. This is not a problem, though, because at the end you will always get the same result.

EXERCISES

1. Perform the GJ elimination:

a) $\begin{bmatrix} 1 & 2 & 7 \\ 5 & 8 & 1 \end{bmatrix}$;

b) $\begin{bmatrix} 1 & 2 & 7 \\ 2 & 4 & 14 \end{bmatrix}$;

c) $\begin{bmatrix} 0 & 2 & 6 \\ 2 & 4 & 11 \end{bmatrix}$;

d) $\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 6 & 10 \\ 0 & 3 & 2 & 7 \end{bmatrix}$;

e) $\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 6 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}$;

f) $\begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix}$;

g) $\begin{bmatrix} 3 & 5 & 2 & 2 \end{bmatrix}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. An example of elimination. Pivot elements marked with square brackets, columns of the identity matrix with round brackets.

a)

$$\begin{bmatrix}[1] & 2 & 7\\ 5 & 8 & 1\end{bmatrix}\quad \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 5w_1 \end{matrix} \rightarrow$$

$$\rightarrow \begin{bmatrix}1 & 2 & 7\\ 0 & [-2] & -34\end{bmatrix}\quad \begin{matrix} w_1' = w_1 + w_2 \\ w_2' = -\frac{1}{2}w_2 \end{matrix} \rightarrow \begin{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} & \begin{pmatrix}0\\1\end{pmatrix} & -27\\ & & 17\end{bmatrix};$$

b)

$$\begin{bmatrix}[1] & 2 & 7\\ 2 & 4 & 14\end{bmatrix}\quad \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \end{matrix} \rightarrow \begin{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} & 2 & 7\\ & 0 & 0\end{bmatrix};$$

c)

$$\begin{bmatrix}0 & 2 & 6\\ 2 & 4 & 11\end{bmatrix}\quad w_1 \leftrightarrow w_2 \quad \rightarrow \begin{bmatrix}[2] & 4 & 11\\ 0 & 2 & 6\end{bmatrix}\quad \begin{matrix} w_1' = \frac{1}{2}w_1 \\ w_2' = w_2 \end{matrix} \rightarrow$$

$$\rightarrow \begin{bmatrix}1 & 2 & \frac{11}{2}\\ 0 & [2] & 6\end{bmatrix}\quad \begin{matrix} w_1' = w_1 - w_2 \\ w_2' = \frac{1}{2}w_2 \end{matrix} \rightarrow \begin{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} & \begin{pmatrix}0\\1\end{pmatrix} & -\frac{1}{2}\\ & & 3\end{bmatrix};$$

d)

$$\begin{bmatrix}[1] & 2 & 3 & 5\\ 2 & 4 & 6 & 10\\ 0 & 3 & 2 & 7\end{bmatrix}\quad \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \\ w_3' = w_3 \end{matrix} \rightarrow \begin{bmatrix}1 & 2 & 3 & 5\\ 0 & 0 & 0 & 0\\ 0 & 3 & 2 & 7\end{bmatrix}\quad w_2 \leftrightarrow w_3 \rightarrow$$

$$\rightarrow \begin{bmatrix}1 & 2 & 3 & 5\\ 0 & [3] & 2 & 7\\ 0 & 0 & 0 & 0\end{bmatrix}\quad \begin{matrix} w_1' = w_1 - \frac{2}{3}w_2 \\ w_2' = \frac{1}{3}w_2 \\ w_3' = w_3 \end{matrix} \rightarrow \begin{bmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} & \begin{pmatrix}0\\1\\0\end{pmatrix} & \frac{5}{3} & \frac{1}{3}\\ & & \frac{2}{3} & \frac{7}{3}\\ & & 0 & 0\end{bmatrix};$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

e)

$$\begin{bmatrix} [1] & 2 & 2 & 1 \\ 2 & 4 & 6 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \\ w_3' = w_3 - 3w_1 \end{matrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & [2] & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix} \begin{matrix} w_1' = w_1 - w_2 \\ w_2' = \frac{1}{2} w_2 \\ w_3' = w_3 - \frac{1}{2} w_2 \end{matrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 0 & -3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & [3] \end{bmatrix} \begin{matrix} w_1' = w_1 + w_3 \\ w_2' = w_2 - \frac{2}{3} w_3 \\ w_3' = \frac{1}{3} w_3 \end{matrix} \rightarrow$$

$$\rightarrow \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad 2 \quad 0 \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right];$$

f)

$$\begin{bmatrix} [2] \\ 5 \\ 7 \end{bmatrix} \begin{matrix} w_1' = \frac{1}{2} w_1 \\ w_2' = w_2 - \frac{5}{2} w_1 \\ w_3' = w_3 - \frac{7}{2} w_1 \end{matrix} \rightarrow \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right];$$

g) $[[3] \quad 5 \quad 2 \quad 2] \quad w_1' = \frac{1}{3} w_1 \rightarrow \left[(1) \quad \frac{5}{3} \quad \frac{2}{3} \quad \frac{2}{3} \right].$

4.3. LINEAR INDEPENDENCE OF VECTORS AND MATRIX RANK

THEORY IN A NUTSHELL

Definition 4.1. Linear combination of vectors
We say that vector z is a linear combination of vectors $x^{(1)}$, $x^{(2)}$, …, $x^{(m)}$, when there exist scalars c_1, c_2, …, c_m such that $z = c_1x^{(1)} + c_2x^{(2)} + \ldots + c_mx^{(m)}$.

For example $z = (1, 2, 4)$ is a linear combination of vectors $x = (2, 0, 10)$ and $y = (0, –1, 1/2)$, because $z = 1/2x – 2y$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Definition 4.2. Linearly independent vectors, linearly dependent vectors
Vectors $x^{(1)}$, $x^{(2)}$, …, $x^{(m)}$ are linearly independent if the zero vector **0** (vector whose all elements are equal to 0) is their linear combination if and only if $c_1 = c_2 = \ldots = c_m = 0$. Otherwise these vectors are linearly dependent.

Linear independence of vectors is related to the **matrix rank**, which you will meet for example while solving the linear systems.

Definition 4.3. Matrix rank
The rank of a matrix A, denoted by $r(A)$, is the maximum number of linearly independent rows or columns that can be found in A.

Determining the matrix rank by definition, would be terribly time-consuming: you would need to examine each combination of rows or each combination of columns in order to check whether their linear combination gives only the zero vector with zero coefficients. Fortunately, there is a much simpler way. To determine the rank of a matrix, you just have to reduce it using **elementary operations** to get the **reduced row echelon form**. The rank of the matrix is equal to the size of the resulting identity matrix.
In order to determine the matrix rank using **WolframAlpha®**, you have to use the command "rank". In particular, if you want to find the rank of the matrix from Example 1 below, then you have to use the following code:

"rank {{1,4,5},{2,10,11},{1,2,4}}".

EXAMPLES

Example 1
Find the rank of the following matrix:

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 10 & 11 \\ 1 & 2 & 4 \end{bmatrix}.$$

Solution
We reduce the matrix, as it has been presented below:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\begin{bmatrix} [1] & 4 & 5 \\ 2 & 10 & 11 \\ 1 & 2 & 4 \end{bmatrix} \begin{matrix} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \\ w_3' = w_3 - w_1 \end{matrix} \rightarrow \begin{bmatrix} 1 & 4 & 5 \\ 0 & [2] & 1 \\ 0 & -2 & -1 \end{bmatrix} \begin{matrix} w_1' = w_1 - 2w_2 \\ w_2' = \frac{1}{2} w_2 \\ w_3' = w_3 + w_2 \end{matrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} & \begin{matrix} 3 \\ \frac{1}{2} \\ 0 \end{matrix} \end{bmatrix}.$$

We managed to obtain two columns of the identity matrix. It means that the rank of A is 2: r(A) = 2.

Example 2
Find the rank of the following matrix:

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Solution
As all the elements of C are 0, it is impossible to reduce any column to the form of a column of the unity matrix. It means that the rank of C is 0: $r(C) = 0$.

EXERCISES

1. Find the ranks of the following matrices:

a) $\begin{bmatrix} 2 & 4 & 14 \\ -5 & -8 & -1 \end{bmatrix}$; b) $\begin{bmatrix} 2 & 4 & 14 \\ 3 & 6 & 21 \end{bmatrix}$;

c) $\begin{bmatrix} 0 & 1 & 3 \\ 1 & 2 & 10 \end{bmatrix}$; d) $\begin{bmatrix} 2 & 4 & 6 & 10 \\ 3 & 6 & 9 & 15 \\ 0 & 6 & 4 & 14 \end{bmatrix}$;

e) $\begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 1 & 3 & 3 \\ 6 & 12 & 14 & 16 \end{bmatrix}$; f) $\begin{bmatrix} 3 \\ 6 \\ 1 \end{bmatrix}$;

g) $\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. a) 2; b) 1; c) 2; d) 2;
   e) 3; f) 1; g) 0.

4.4. DETERMINANTS

THEORY IN A NUTSHELL

Definition 4.5. Cofactor
Cofactor A_{ij} of the element a_{ij} of matrix A is the determinant of the submatrix of A obtained by removing row i and column j, multiplied by $(-1)^{i+j}$.

Definition 4.6. Determinant
The determinant of a **square** matrix A, denoted by $\det(A)$, is a matrix invariant defined as follows. In the case of a matrix of size 1×1 of the form $[a]$, $\det(A) = a$. In the case of a matrix of size $n \times n$, where $n > 1$, the determinant is defined using the cofactors of the elements of any row i or column j:

$$\det(A) = \sum_{j=1}^{n} a_{ij}A_{ij} = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} ,$$

$$\det(A) = \sum_{i=1}^{n} a_{ij}A_{ij} = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} .$$

As you can see, in order to find the determinant, you must add the products of all the elements of a chosen row and their cofactors, or do the same with any column. Calculating determinants by definition takes much time, so it is important to remember their properties.

Theorem 4.1. Properties of determinants

1. The determinant of a matrix of size 2×2

   $$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

   equals to:

   $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

(we reduce the product of the elements of the main diagonal by the product of the elements of the other diagonal).

2. The determinant of a matrix of size 3×3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

equals to[10]:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

3. If all elements of any row or of any column of a matrix are 0, then its determinant equals to 0.
4. If in any row or any column there is only one nonzero element, say a_{ij}, then $\det(A) = a_{ij}A_{ij}$.
5. Performing the elementary operation of subtracting a multiplicity of one row from another row does not change the value of the determinant.
6. Performing the elementary operation of multiplying a row by a number c causes multiplication of the determinant by c. In other words, if A' is the matrix obtained from A by multiplying any row of A by c, then $\det(A') = c \cdot \det(A)$, or equivalently $\det(A) = 1/c \cdot \det(A')$.
7. Performing the elementary operation of swapping two rows causes multiplication of the determinant by –1. In other words, if A' is the matrix obtained from A by swapping any two rows, then $\det(A') = (-1) \cdot \det(A)$, or equivalently $\det(A) = (-1) \cdot \det(A')$.

In order to determine the determinant using **WolframAlpha®**, you can use the command "det". In particular, if you want to find the determinant of the matrix from Example 1 below, then you have to use the following code:

"det {{2,3},{5,7}}".

[10] We proceed similarly as in the case of a matrix of size 2×2. First we repeat the two first columns writing it on the right (or first two rows below the matrix). Then we add the products of the elements of each upper left – bottom right diagonal and subtract the products of the elements of each upper right – bottom left diagonal. **Note**: this method does not work for larger matrices!

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1
Find the determinant of the following matrix:

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

Solution
As A has size 2×2, we use the property 1: $\det(A) = 2 \cdot 7 - 3 \cdot 5 = -1$.

Example 2
Find the determinant of the following matrix:

$$A = \begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 8 \\ 3 & 4 & 3 \end{bmatrix}.$$

Solution
This time the size of A is 3×3, so we use the property 2. We begin with copying the two first columns:

$$\begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 8 \\ 3 & 4 & 3 \end{bmatrix} \begin{matrix} 3 & 1 \\ 2 & 1 \\ 3 & 4 \end{matrix}$$

Now we add the products of the elements of each upper left – bottom right diagonal $3 \cdot 1 \cdot 3$, $1 \cdot 8 \cdot 3$ i $4 \cdot 2 \cdot 4$, and subtract the products of the elements of each upper right – bottom left diagonal: $4 \cdot 1 \cdot 3$, $3 \cdot 8 \cdot 4$ i $1 \cdot 2 \cdot 3$. Finally we obtain:

$$\begin{aligned} \det(A) &= 3 \cdot 1 \cdot 3 + 1 \cdot 8 \cdot 3 + 4 \cdot 2 \cdot 4 - 4 \cdot 1 \cdot 3 - 3 \cdot 8 \cdot 4 - 1 \cdot 2 \cdot 3 = \\ &= 9 + 24 + 32 - 12 - 96 - 6 = -49. \end{aligned}$$

Example 3
Find the determinant of the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 \\ 1 & 5 & 4 & 3 \\ 3 & -2 & -3 & 1 \\ 1 & 2 & 5 & 2 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

The size of A is greater than 3×3, so we have to use the elementary operations. We first reduce the column 1:

$$\begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - w_1 \\ w_3' = w_3 - 3w_1 \\ w_4' = w_4 - w_1 \end{array} \rightarrow \begin{bmatrix} 1 & 1 & 3 & 1 \\ 0 & 4 & 1 & 2 \\ 0 & -5 & -12 & -2 \\ 0 & 1 & 2 & 1 \end{bmatrix}.$$

It follows from the properties 5 and 6, that the determinant of new matrix is equal to the determinant of A. As in the first column we have now exactly one nonzero element (the element [1, 1]), we can use the property 4:

$$\det(A) = (-1)^{1+1} \cdot 1 \cdot \det\left(\begin{bmatrix} 4 & 1 & 2 \\ -5 & -12 & -2 \\ 1 & 2 & 1 \end{bmatrix}\right).$$

Using the property 2, we finally obtain $\det(A) = -25$.

EXERCISES

1. Find the determinants of the following matrices:

a) $\begin{bmatrix} 4 \end{bmatrix}$;

b) $\begin{bmatrix} 2 & -2 \\ 1 & 3 \end{bmatrix}$;

c) $\begin{bmatrix} 2 & -3 \\ 2 & 4 \end{bmatrix}$;

d) $\begin{bmatrix} 1 & 5 & 7 \\ 2 & -3 & 2 \\ 1 & 0 & 0 \end{bmatrix}$;

e) $\begin{bmatrix} 0 & -2 & -2 \\ 2 & 4 & 2 \\ 3 & 2 & 0 \end{bmatrix}$;

f) $\begin{bmatrix} 3 & 3 & 3 \\ 2 & 2 & 2 \\ 1 & 1 & 1 \end{bmatrix}$;

g) $\begin{bmatrix} 2 & 2 & 1 & 4 \\ 5 & 6 & 8 & 1 \\ 2 & 3 & 1 & 1 \\ 4 & 4 & 2 & 8 \end{bmatrix}$;

h) $\begin{bmatrix} 0 & 2 & 3 & 1 \\ 0 & 3 & 3 & 2 \\ 2 & -7 & -2 & -5 \\ 4 & 11 & 2 & 1 \end{bmatrix}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

i) $\begin{bmatrix} 1 & 2 & 7 & 2 & 3 \\ 5 & 1 & 7 & 2 & -4 \\ -3 & 2 & 7 & 1 & 4 \\ 5 & -2 & -2 & 3 & 5 \\ 1 & 2 & 2 & 4 & 1 \end{bmatrix}$.

SOLUTIONS

1. a) 4; b) 8; c) 14;
   d) 31; e) 4; f) 0;
   g) 0; h) 72; i) 1653.

4.5. MATRIX INVERSE

THEORY IN A NUTSHELL

As you probably remember, the matrix multiplication is not commutative in general case – AB is usually not the same as BA. One of the exceptions is multiplication by the inverse (if it exists).

Definition 4.7. Matrix inverse
Assume that A is a square matrix. The **inverse** of A is the matrix A^{-1} satisfying the condition $AA^{-1} = A^{-1}A = I$, where I is identity matrix of same size as A.

Deriving the matrix inverse from definition takes much time. That's why you will use one of the two following methods:

Method 1
We augment A by writing the identity matrix I on its right. We reduce the augmented matrix with the elementary operations in such a way that in the end the identity matrix appears on the left. If it is possible, then the matrix appearing on the right is exactly A^{-1}.

Method 2

We use the formula: $A^{-1} = \frac{A_c^T}{\det(A)}$. The matrix in the numerator is the **adjugate matrix**, being the transpose of the **matrix of cofactors** (i.e., the matrix in which each element is substituted with its cofactor).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that in the second method we divide the adjugate matrix by the determinant of A. If $\det(A) = 0$, this operation is impossible. We then say that A is **singular**, which means that it does not have an inverse. If you use the first method, in a singular matrix a row of zeros will appear in some step.
The second method is extremely useful in the case of the matrices of size 2×2, because it reduces to the following simple formula:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}}{a_{11}a_{22} - a_{12}a_{21}}.$$

(we swap the elements on the main diagonal, we change signs of the two remaining elements and divide the result by the determinant).
In order to find the matrix inverse using **WolframAlpha®**, you can use the command "inv", or the symbol of power "^". In particular, if you want to inverse the matrix from Example 1 below, then you can use the following code:

"inv {{1,3},{-2,4}}",

or this one:

"{{1,3},{-2,4}}^(-1)".

EXAMPLES

Example 1
Find the inverse of the matrix:

$$A = \begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix}.$$

Solution
We used the simplified formula:

$$A^{-1} = \frac{\begin{bmatrix} 4 & -3 \\ 2 & 1 \end{bmatrix}}{1 \cdot 4 - 3 \cdot (-2)} = \begin{bmatrix} \frac{2}{5} & -\frac{3}{10} \\ \frac{1}{5} & \frac{1}{10} \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 2
Find the inverse of the following matrix using two methods:

$$A=\begin{bmatrix} 1 & -2 & 4 \\ -1 & 3 & -6 \\ 2 & -1 & 3 \end{bmatrix}.$$

Solution
Method 1
We begin with augmenting the matrix:

$$\left[\begin{array}{ccc|ccc} 1 & -2 & 4 & 1 & 0 & 0 \\ -1 & 3 & -6 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 & 0 & 1 \end{array}\right].$$

Now we perform the elimination (remember that we act on whole, augmented matrix):

$$\left[\begin{array}{ccc|ccc} 1 & -2 & 4 & 1 & 0 & 0 \\ -1 & 3 & -6 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 + w_1 \\ w_3' = w_3 - 2w_1 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & -2 & 4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & 0 \\ 0 & 3 & -5 & -2 & 0 & 1 \end{array}\right] \begin{array}{l} w_1' = w_1 + 2w_2 \\ w_2' = w_2 \\ w_3' = w_3 - 3w_2 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 0 \\ 0 & 1 & -2 & 1 & 1 & 0 \\ 0 & 0 & 1 & -5 & -3 & 1 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 + 2w_3 \\ w_3' = w_3 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 0 \\ 0 & 1 & 0 & -9 & -5 & 2 \\ 0 & 0 & 1 & -5 & -3 & 1 \end{array}\right].$$

This means that:

$$A^{-1}=\begin{bmatrix} 3 & 2 & 0 \\ -9 & -5 & 2 \\ -5 & -3 & 1 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Method 2

Our first step is finding the determinant of A: $\det(A) = 1$.

Now we are going to find the cofactors of all the elements of A. Recall that the cofactor of an element a_{ij} (i.e., the element in row i and column j) is equal to the determinant of the submatrix obtained by removing the row i and column j from A, multiplied by $(-1)^{i+j}$. Thus we obtain:

$$A_{11} = (-1)^{1+1} \det\left(\begin{bmatrix} 3 & -6 \\ -1 & 3 \end{bmatrix}\right) = 3,$$

$$A_{12} = (-1)^{1+2} \det\left(\begin{bmatrix} -1 & -6 \\ 2 & 3 \end{bmatrix}\right) = -9,$$

$$A_{13} = (-1)^{1+3} \det\left(\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix}\right) = -5,$$

$$A_{21} = (-1)^{2+1} \det\left(\begin{bmatrix} -2 & 4 \\ -1 & 3 \end{bmatrix}\right) = 2,$$

$$A_{22} = (-1)^{2+2} \det\left(\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}\right) = -5,$$

$$A_{23} = (-1)^{2+3} \det\left(\begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix}\right) = -3,$$

$$A_{31} = (-1)^{3+1} \det\left(\begin{bmatrix} -2 & 4 \\ 3 & -6 \end{bmatrix}\right) = 0,$$

$$A_{32} = (-1)^{3+2} \det\left(\begin{bmatrix} 1 & 4 \\ -1 & -6 \end{bmatrix}\right) = 2,$$

$$A_{33} = (-1)^{3+3} \det\left(\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}\right) = 1.$$

By inserting those values into the matrix, we obtain the **cofactor matrix**:

$$A_c = \begin{bmatrix} 3 & -9 & -5 \\ 2 & -5 & -3 \\ 0 & 2 & 1 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

After transposing it and using the appropriate formula, we finally obtain

$$A^{-1} = \begin{bmatrix} 3 & 2 & 0 \\ -9 & -5 & 2 \\ -5 & -3 & 1 \end{bmatrix} / 1 = \begin{bmatrix} 3 & 2 & 0 \\ -9 & -5 & 2 \\ -5 & -3 & 1 \end{bmatrix}.$$

Example 3
Find the inverse of the following matrix using two methods:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 2 & -1 & 5 \end{bmatrix}.$$

Solution
Method 1
The augmented matrix and the elementary operations have been presented below:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 2 & 4 & 8 & 0 & 1 & 0 \\ 2 & -1 & 5 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \\ w_3' = w_3 - 2w_1 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 & 0 \\ 0 & -5 & -3 & -2 & 0 & 1 \end{array}\right].$$

As you can see, the second row is filled with 0 (we mean the matrix on the left). It means that A is singular (the inverse does not exist).

Method 2
We begin with finding the determinant of A. It is equal to 0, so A is singular.

EXERCISES

1. Find the inverses, using arbitrary method:

   a) $[7]$; b) $\begin{bmatrix} 1 & 5 \\ 2 & -4 \end{bmatrix}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) $\begin{bmatrix} 2 & 3 \\ -1 & -\frac{3}{2} \end{bmatrix}$;

d) $\begin{bmatrix} 3 & -4 & 2 \\ -2 & 3 & -3 \\ 2 & 1 & 3 \end{bmatrix}$;

e) $\begin{bmatrix} 2 & 5 & 7 \\ -2 & 1 & -3 \\ 1 & 1 & 3 \end{bmatrix}$;

f) $\begin{bmatrix} 4 & 2 & 1 \\ 3 & -1 & 5 \\ 5 & 5 & -3 \end{bmatrix}$;

g) $\begin{bmatrix} 1 & 2 & 10 & 2 \\ 5 & 1 & 7 & 2 \\ -3 & 2 & 8 & 1 \\ 0 & -2 & -2 & 3 \end{bmatrix}$;

h) $\begin{bmatrix} 3 & 5 & -8 & 2 \\ 2 & -1 & -3 & 3 \\ -2 & 5 & -1 & 8 \\ -3 & -1 & 4 & 9 \end{bmatrix}$.

SOLUTIONS

1. a) $\begin{bmatrix} \frac{1}{7} \end{bmatrix}$; b) $\begin{bmatrix} \frac{2}{7} & \frac{5}{14} \\ \frac{1}{7} & -\frac{1}{14} \end{bmatrix}$; c) singular; d) $\begin{bmatrix} \frac{3}{5} & \frac{7}{10} & \frac{3}{10} \\ 0 & \frac{1}{4} & \frac{1}{4} \\ -\frac{2}{5} & -\frac{11}{20} & \frac{1}{20} \end{bmatrix}$;

e) $\begin{bmatrix} 1 & -\frac{4}{3} & -\frac{11}{3} \\ \frac{1}{2} & -\frac{1}{6} & -\frac{4}{3} \\ -\frac{1}{2} & \frac{1}{2} & 2 \end{bmatrix}$; f) singular; g) $\begin{bmatrix} -3 & 2 & 2 & 0 \\ -\frac{205}{2} & 64 & \frac{145}{2} & \frac{3}{2} \\ \frac{61}{2} & -19 & -\frac{43}{2} & -\frac{1}{2} \\ -48 & 30 & 34 & 1 \end{bmatrix}$;

h) singular.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 5
LINEAR SYSTEMS

5.1. SYSTEMS OF LINEAR EQUATIONS

THEORY IN A NUTSHELL

You probably know what a linear equation or system of such equations, and therefore we omit their definitions. Important, however, is the notation which we will use. The coefficients of the variables form the **matrix of the system** A. The right side coefficients form the **vector of intercepts** b, also called the **right hand side**. For example, if the system of linear equations has the form:

$$\begin{cases} 2x_1 + 3x_2 - x_3 = 7, \\ x_1 + x_2 + 2x_3 = 5, \end{cases}$$

then the above mentioned matrices are:

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 1 & 2 \end{bmatrix} \text{ i } b = \begin{bmatrix} 7 \\ 5 \end{bmatrix},$$

and the variables are x_1, x_2 and x_3. **A solution of the system** are such values of the variables (often denoted by a vector), which satisfy the system when we substitute them to the variables. In the example one of solutions is $x_1 = 1$, $x_2 = 2$, $x_3 = 1$ (we can also write it as $x = (1, 2, 1)$). It can happen that a system has one solution (we say then that it is **fully determined**), it can have infinitely many solutions (**underdetermined** system; for example the above system is underdetermined – now you do not see that but soon you will be able to check it). Finally, it is possible that a system does not have any solution. In such a case we say that it is **inconsistent**.
In order to solve a given system, we write it in the matrix form, together with the right hand side. This way we construct the **augmented matrix**. The augmented matrix of the above system has the form (note that there is a vertical line separating its two parts):

$$\left[\begin{array}{ccc|c} 2 & 3 & -1 & 7 \\ 1 & 1 & 2 & 5 \end{array}\right].$$

Then we reduce the augmented matrix (note that we reduce only the columns of A, i.e. the columns of the left of the vertical line). If at some step there appears

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

a row where the intercept is different from zero, and all other elements are zero, then the system is inconsistent. Otherwise, we read the reduced form of the system from the matrix. If each variable is assigned a specific value, we have a fully determined system whose reduced form is its only solution. Otherwise (i.e., when in at least one of the equations, there are at least two variables), we are dealing with a underdetermined system. In this case, we find in the reduced form of matrix *A* the columns that **do not form** the identity submatrix and treat them as parameters. With the parameters we can express all the variables. This way we obtain the general solution. In case of an underdetermined system very important are also the so-called **basic solutions** defined as follows. A **base** is any set of variables of cardinality equal to the rank of *A*, containing all the variables whose values are fixed (i.e. variables expressed without parameters in the reduced form of the system). The members of a base are called **basic variables**, the remaining ones are **non-basic variables**. All the non-basic variables take the value 0, and then the values of basic variables are derived from the reduced system.
In order to solve a system of linear equations using **WolframAlpha®**, you can write the equations, separating them with commas (","), where the begin and end of the system are marked with curly brackets "{" and "}". The whole statement should begin with the word "solve". In particular, in order to solve the system from Example 1 below, you can type:

"solve {x1+2x2=8, 2x1+3x2=13}".

You can also decide, which variables should be expressed with the other ones by ending the statement with the word "for" followed by a list of variables in curly brackets. For example, in order to find a general solution of the system from Example 3 below, where x_1 and x_2 are expressed by x_3, you can type:

"solve {x1+2x2+x3=8, 2x1+x2+5x3=20} for {x1,x2}".

EXAMPLES

Example 1
Solve the following system of equations:

$$\begin{cases} x_1 + 2x_2 = 8, \\ 2x_1 + 3x_2 = 13. \end{cases}$$

Solution
We begin with writing the system in the matrix form. Then we reduce it, using the elementary operations:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\left[\begin{array}{cc|c} 1 & 2 & 8 \\ 2 & 3 & 13 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \end{array} \rightarrow \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & -1 & -3 \end{array}\right] \begin{array}{l} w_1' = w_1 + 2w_2 \\ w_2' = -w_2 \end{array} \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \end{array}\right].$$

From the last matrix we obtain:

$$\begin{cases} 1x_1 + 0x_2 = 2, \\ 0x_1 + 1x_2 = 3, \end{cases}$$

or simply:

$$\begin{cases} x_1 = 2, \\ x_2 = 3. \end{cases}$$

It means that the system has exactly one solution $x_1 = 2$, $x_2 = 3$, which can be also written as $x = (2, 3)$. It is a fully-determined system.

Example 2
Solve the following system of equations:

$$\begin{cases} 2x_1 + x_2 - 4x_3 = 1, \\ 2x_1 - 3x_3 = 2, \\ 2x_1 + 2x_2 - 5x_3 = 5. \end{cases}$$

Solution
We begin with writing the system in the matrix form. Note that the variable x_2 is not present in the second equation and thus the element [2, 2] is equal to 0. Then we perform the elementary operations:

$$\left[\begin{array}{ccc|c} 2 & 1 & -4 & 1 \\ 2 & 0 & -3 & 2 \\ 2 & 2 & -5 & 5 \end{array}\right] \begin{array}{l} w_1' = \frac{1}{2} w_1 \\ w_2' = w_2 - w_1 \\ w_3' = w_3 - w_1 \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & -2 & \frac{1}{2} \\ 0 & -1 & 1 & 1 \\ 0 & 1 & -1 & 4 \end{array}\right] \begin{array}{l} w_1' = w_1 + \frac{1}{2} w_2 \\ w_2' = -w_2 \\ w_3' = w_3 + w_2 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{3}{2} & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 5 \end{array}\right].$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

As you can see, the intercept in the third row is not equal to 0 (it is equal to 5), but all the other elements of this row are 0. It means that the system is inconsistent – it does not have a solution (if we wrote the third equation it would take the form $0x_1 + 0x_2 + 0x_3 = 5$, i.e. $0 = 5$, a contradiction).

Example 3
Solve the following system of equations:

$$\begin{cases} x_1 + 2x_2 + x_3 = 8, \\ 2x_1 + x_2 + 5x_3 = 20. \end{cases}$$

Solution
The elimination of the augmented matrix is as follows:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 2 & 1 & 5 & 20 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & -3 & 3 & 4 \end{array}\right] \begin{array}{l} w_1' = w_1 + \frac{2}{3} w_2 \\ w_2' = -\frac{1}{3} w_2 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 3 & \frac{32}{3} \\ 0 & 1 & -1 & -\frac{4}{3} \end{array}\right].$$

We can read the following reduced form of the system:

$$\begin{cases} x_1 + 3x_3 = \frac{32}{3}, \\ x_2 - x_3 = -\frac{4}{3}. \end{cases}$$

As we can see, the identity submatrix is formed by the two first columns, i.e. the columns of variables x_1 and x_2. It means that the remaining variables (in this case only one variable, x_3) are parameters. We denote the parameter by a, where $a \in R$. By substituting it to the equations we finally obtain:

$$\begin{cases} x_1 = \frac{32}{3} - 3a, \\ x_2 = -\frac{4}{3} + a, \\ x_3 = a, \qquad a \in R. \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

It can be written in an equivalent form $x = (32/3 - 3a, -4/3 + a, a)$, $a \in R$. It is the general solution of the system which turned out to be underdetermined.
In such a situation we should determine all the basic solutions. To do this, first we print all sets of variables of cardinality equal to the rank of the matrix of the system. Note that we do not need again to reduce the matrix to determine its rank: we already know its reduced form and we can conclude that $r(A) = 2$. So we have to write all the combinations of two variables, bearing in mind that the variables with fixed values must be present in such combinations. In this example, there are no such variables (in the general solution, all variables are represented using a parameter). The combinations of variables are then $\{x_1, x_2\}$, $\{x_1, x_3\}$ and $\{x_2, x_3\}$. Taking them as bases, we can find the basic solutions. Let us start with the base $B^{(1)} = \{x_1, x_2\}$[11]. The basic variables are x_1 and x_2, so the non-basic variable is x_3. Hence $x_3 = 0$. We substitute this value to the reduced system:

$$\begin{cases} x_1 + 3 \cdot 0 = \dfrac{32}{3}, \\ x_2 - 0 = -\dfrac{4}{3}. \end{cases}$$

Which imples that $x_1 = 32/3$ and $x_2 = -4/3$, so the first basic solution is $x^{(1)} = (32/3, -4/3, 0)$. In case of the second base, $B^{(2)} = \{x_1, x_3\}$, the basic variables are x_1 and x_3, so the non-basic variable is x_2. Hence $x_2 = 0$. Substituting it to the reduced system:

$$\begin{cases} x_1 + 3x_3 = \dfrac{32}{3}, \\ 0 - x_3 = -\dfrac{4}{3}. \end{cases}$$

We obtain $x_1 = 20/3$ and $x_3 = 4/3$, which gives us the second basic solution $x^{(2)} = (20/3, 0, 4/3)$. Finally, the case of the third base $B^{(3)} = \{x_2, x_3\}$ the basic variables are x_2 and x_3, so the non-basic variable is x_1. Hence $x_1 = 0$. By substituting this value to the reduced system:

$$\begin{cases} 0 + 3x_3 = \dfrac{32}{3}, \\ x_2 - x_3 = -\dfrac{4}{3}. \end{cases}$$

[11] The order and method of enumeration of the bases and basic solutions does not matter, of course.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

we obtain $x_2 = 20/9$ and $x_3 = 32/9$, which allows us to find the last basic solution $x^{(3)} = (0, 20/9, 32/9)$. To conclude, the system has three basic solutions: $x^{(1)} = (32/3, -4/3, 0)$, $x^{(2)} = (20/3, 0, 4/3)$ and $x^{(3)} = (0, 20/9, 32/9)$.

Example 4
Solve the following system of equations:

$$\begin{cases} x_1 + 3x_2 + 2x_3 - x_4 = 4, \\ 2x_1 + 6x_2 + 5x_3 - 2x_4 = 10. \end{cases}$$

Solution
The augmented matrix and the elementary operations are as follows:

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & -1 & 4 \\ 2 & 6 & 5 & -2 & 10 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 2 & -1 & 4 \\ 0 & 0 & 1 & 0 & 2 \end{array}\right] \begin{array}{l} w_1' = w_1 - 2w_2 \\ w_2' = w_2 \end{array} \rightarrow \left[\begin{array}{cccc|c} 1 & 3 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{array}\right].$$

The reduced system is:

$$\begin{cases} x_1 + 3x_2 - x_4 = 0, \\ x_3 = 2. \end{cases}$$

In the reduced matrix, the identity matrix is formed by the columns of x_1 and x_3, so x_2 and x_4 are parameters (we denote them by a_1 and a_2, respectively). This way we obtain the following form of the general solution: $x_1 = -3a_1 + a_2$, $x_2 = a_1$, $x_3 = 2$, $x_4 = a_2$, where $a_1, a_2 \in R$.
In order to find the basic solutions, we begin with listing all the bases. Observe that the rank of the matrix of system is $r(A) = 2$, and the variable x_3 is fixed ($x_3 = 2$). It means that we are interested only in those two-element sets of variables which contain x_3: $\{x_1, x_3\}$, $\{x_2, x_3\}$ and $\{x_3, x_4\}$. We do not take into account the remaining combinations of two variables, i.e. $\{x_1, x_2\}$, $\{x_1, x_4\}$ or $\{x_2, x_4\}$, because they do not contain x_3. In the case of $B^{(1)} = \{x_1, x_3\}$ the basic variables are x_1 and x_3, and the non-basic variables x_2 and x_4. By substituting $x_2 = 0$ and $x_4 = 0$ to the reduced system, we obtain $x^{(1)} = (0, 0, 2, 0)$. In the case of $B^{(2)} = \{x_2, x_3\}$ the basic variables are x_2 and x_3, and the non-basic variables are x_1 and x_4. By substituting $x_1 = 0$ and $x_4 = 0$ to the reduced system we obtain $x^{(2)} = (0, 0, 2, 0)$. In the case of $B^{(3)} = \{x_3, x_4\}$ the basic variables are x_3 and x_4, and the non-basic variables are x_1 and x_2. By substituting $x_1 = 0$ and $x_2 = 0$ to the reduced system, we obtain $x^{(3)} = (0, 0, 2, 0)$. Note that all the three basic solutions are in fact the same. It follows from the fact that a basic variable is

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

equal to 0 in each of them (x_1 in the first solution, x_2 in the second one, and x_4 in the third one). We call such a situation **degeneracy** and such solutions are called **degenerate solutions**.

EXERCISES

1. Solve the following systems of equations:

a) $\begin{cases} 3x_1 + x_2 = 7, \\ x_1 - x_2 = 1; \end{cases}$

b) $\begin{cases} 2x_1 - 3x_2 + 2x_3 = 2, \\ 3x_1 + x_2 + x_3 = 12, \\ x_1 + 4x_2 + x_3 = 12; \end{cases}$

c) $\begin{cases} 7x_1 + 11x_2 = 20, \\ 14x_1 + 22x_2 = 50; \end{cases}$

d) $\begin{cases} 2x_1 - 4x_2 + 2x_3 = 10, \\ 3x_1 + 2x_2 + 7x_3 = 12, \\ x_1 - 10x_2 - 3x_3 = 15; \end{cases}$

e) $\begin{cases} x_1 + 2x_2 = 40, \\ 3x_1 + 6x_2 = 120; \end{cases}$

f) $\begin{cases} x_1 + 2x_2 - 7x_3 = 10, \\ 3x_1 + 8x_2 + 5x_3 = 6; \end{cases}$

g) $\begin{cases} 2x_1 + 2x_2 - 7x_3 + 5x_4 = 6, \\ 2x_1 + 2x_2 + 5x_4 = 13; \end{cases}$

h) $\begin{cases} x_1 + 3x_2 - 4x_3 + x_4 = 1, \\ 3x_1 + 8x_2 - 7x_3 + 2x_4 = 6, \\ x_1 + 4x_2 - 9x_3 + 2x_4 = -2; \end{cases}$

i) $\begin{cases} x_1 + 2x_2 - 5x_3 + 2x_4 = 4, \\ 2x_1 + 5x_2 - 10x_3 + 4x_4 = 10, \\ 3x_1 + 7x_2 - 15x_3 + 7x_4 = 15; \end{cases}$

j) $\begin{cases} 2x_1 + x_2 + 3x_3 - 7x_4 = -1, \\ 3x_1 + 2x_2 + 7x_3 - 4x_4 = 6, \\ x_1 - x_3 + 5x_4 = 7. \end{cases}$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. Sample general solutions and all basic solutions (b.s.):
   a) $x_1 = 2, x_2 = 1$;
   b) $x_1 = 3, x_2 = 2, x_3 = 1$;
   c) inconsistent;
   d) inconsistent;
   e) $x_1 = 40 - 2a, x_2 = a$, where $a \in R$, b.s.: (40, 0), (0, 20);
   f) $x_1 = 34 + 33a, x_2 = -12 - 13a, x_3 = a$, where $a \in R$;
      b.s.: (34, –12, 0), (46/13, 0, –12/13), (0, 46/33, –34/33);
   g) $x_1 = 13/2 - a_1 - 5/2a_2,\ x_2 = a_1,\ x_3 = 1,\ x_4 = a_2,$ where $a_1, a_2 \in R$;
      b.s.: (13/2, 0, 1, 0), (0, 13/2, 1, 0), (0, 0, 1, 13/5);
   h) $x_1 = 10 - 11a_1 + 2a_2,\ x_2 = -3 + 5a_1 - a_2,\ x_3 = a_1,\ x_4 = a_2,$ where $a_1, a_2 \in R$; b.s.: (10, –3, 0, 0), (17/5, 0, 3/5, 0), (4, 0, 0, –3), (0, 17/11, 10/11, 0), (0, 2, 0, –5), (0, 0, 4, 17);
   i) $x_1 = -2 + 5a,\ x_2 = 2,\ x_3 = a,\ x_4 = 1,$ where $a \in R$;
      b.s.: (–2, 2, 0, 1), (0, 2, 2/5, 1);
   j) $x_1 = 2 + a,\ x_2 = 2 - 5a,\ x_3 = a,\ x_4 = 1,$ where $a \in R$;
      b.s.: (2, 2, 0, 1), (12/5, 0, 2/5, 1), (0, 12 –2, 1).

5.2. SYSTEMS OF LINEAR INEQUALITIES

THEORY IN A NUTSHELL

As in the case of equations, also in the case of linear inequalities we do not give a formal definition of them. Solving systems of inequalities is reduced to solving systems of equations. You only need to modify the left side of each inequality belonging to the system. More specifically, to each inequality will be assigned an additional variable, the so-called **slack variable**.
If an inequality has the form "≤", then we add the slack variable, otherwise we subtract it. For example, the system:

$$\begin{cases} 2x_1 + 3x_2 \le 4, \\ 5x_1 + 6x_2 \ge 7, \end{cases}$$

takes the following form after including the slack variables:

$$\begin{cases} 2x_1 + 3x_2 + x_3 = 4, \\ 5x_1 + 6x_2 - x_4 = 7. \end{cases}$$

Note: the slack variables must be always nonnegative.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The augmented matrix consists of three segments: the first segment is a matrix of coefficients corresponding to the left hand side of the inequalities, the second (the middle one) contains the coefficients of the slack variables (i.e. it is an identity matrix with some elements eventually multiplied by –1), finally the third segment is formed by the intercepts. During the elimination of the matrix the objective is to obtain the reduced form of the first segment. Like in the case of a system of equations, there are various situations. This time, however, we consider only two cases.

If, after elimination of the first segment of the matrix there is no row consisting of zeros, then we can determine the general solution. Slack variables are always parameters. Note that the parameters corresponding to the "regular" variable can take any real value (as in the systems of equations), but the parameters corresponding to the slack variables must be nonnegative. In the course of writing the basic solutions you must have in mind to eliminate all those with even one slack variable being negative. Also remember that writing a general solution, we write only the "regular" variables, but in the case of basic solutions – we include the slack variables, too.

If, after elimination of the first segment of the matrix there is at least one row consisting of zeros, we need to check the consistency of the system. For this purpose, we write the remainder of the zero rows (and therefore part of the second and third segments) and the corresponding system of equations. If it has even one basic solution consisting only of nonnegative variables, then initial system of inequalities is consistent and we proceed further as in the first case (you only need to remember the additional conditions on the parameters corresponding to the slack variables; analyze carefully the Example 2). But if every basic solution of the smaller system of equations contains even one negative variable, then the initial system of inequalities is inconsistent. **Note**: You can go directly to the determination of basic solutions of the underlying initial system of equations. If it is contradictory it will take a lot more time. However, this method will be used in the next section – in the case of linear programming problems, especially those corresponding to the real economic problems, a contradiction occurs very rarely.

In order to solve a system of linear inequalities using **WolframAlpha®**, you must first transform it to the form of the system of equations, and then solve it as in the case of ordinary system of equations, adding the nonnegativity assumption for the slack variables. For example, in order to solve the system from Example 1 below, you can type:

"solve {2x1+4x2+x3=12, 2x1+x2+x4=6, x3>=0, x4>=0} for {x1,x2}".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1

Solve the system of inequalities:

$$\begin{cases} 2x_1 + 4x_2 \le 12, \\ 2x_1 + \ x_2 \le \ 6. \end{cases}$$

Solution

We begin with introducing the slack variables: x_3 into the first inequality and x_4 into the second one. As they both have the form "≤", both variables are added to the left hand sides. This way we obtain the system of equations:

$$\begin{cases} 2x_1 + 4x_2 + x_3 \qquad = 12, \\ 2x_1 + \ x_2 \qquad + x_4 = \ 6. \end{cases}$$

where $x_3, x_4 \ge 0$. The augmented matrix is shown below (first matrix); Also the reduction process has been presented. Observe that the second segment is the identity matrix. It will be always like this, because in each inequality there is another slack variable (it can eventually happen that a 1 can be replaced with –1, if the corresponding inequality is of the type "≥").

$$\left[\begin{array}{cc|cc|c} 2 & 4 & 1 & 0 & 12 \\ 2 & 1 & 0 & 1 & 6 \end{array}\right] \begin{array}{l} w_1' = \frac{1}{2} w_1 \\ w_2' = w_2 - w_1 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{cc|cc|c} 1 & 2 & \frac{1}{2} & 0 & 6 \\ 0 & -3 & -1 & 1 & -6 \end{array}\right] \begin{array}{l} w_1' = w_1 + \frac{2}{3} w_2 \\ w_2' = -\frac{1}{3} w_2 \end{array} \rightarrow \left[\begin{array}{cc|cc|c} 1 & 0 & -\frac{1}{6} & \frac{2}{3} & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} & 2 \end{array}\right].$$

There are no zeroed rows in the first segment, so we do not have to verify the consistency (the system is consistent for sure). We can read the following reduced version of the system:

$$\begin{cases} x_1 \qquad - \frac{1}{6} x_3 + \frac{2}{3} x_4 = 2, \\ \qquad x_2 + \frac{1}{3} x_3 - \frac{1}{3} x_4 = 2. \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

As the identity matrix is formed by the columns corresponding with x_1 and x_2, there are no parameters among the "regular" variables. The parameters are only slack variables. We can assume that x_3 and x_4 are b_1 and b_2, respectively and write the general solution:
$x_1 = 2 + 1/6b_1 - 2/3b_2$, $x_2 = 2 - 1/3b_1 + 1/3b_2$, where $b_1, b_2 \in R_+ \cup \{0\}$.
Let us determine the basic solutions. Observe that the rank of the matrix of the system is $r(A) = 2$, so we must check all the combinations of two variables out of four. There are six of them. In the case of the base $B^{(1)} = \{x_1, x_2\}$, we substitute $x_3 = x_4 = 0$ to the reduced system and we obtain $x^{(1)} = (2, 2, 0, 0)$. In the case of the base $B^{(2)} = \{x_1, x_3\}$, we substitute $x_2 = x_4 = 0$ to the reduced system and we obtain $x^{(2)} = (3, 0, 6, 0)$. In the case of the base $B^{(3)} = \{x_1, x_4\}$, we substitute $x_2 = x_3 = 0$ to the reduced system and we obtain $x^{(3)} = (6, 0, 0, -6)$. We skip this solution, as one of the variables is negative ($x_4 = -6$). In the case of the base $B^{(4)} = \{x_2, x_3\}$, we substitute $x_1 = x_4 = 0$ to the reduced system and we obtain $x^{(4)} = (0, 6, -12, 0)$. We reject also this solution ($x_3 = -12 < 0$). In the case of the base $B^{(5)} = \{x_2, x_4\}$, we substitute $x_1 = x_3 = 0$ to the reduced system and we obtain $x^{(5)} = (0, 3, 0, 3)$. In the case of the base $B^{(6)} = \{x_3, x_4\}$, we substitute $x_1 = x_2 = 0$ to the reduced system and we obtain $x^{(6)} = (0, 0, 12, 6)$. Finally we can say that the system has exactly four basic **feasible** solutions: $x^{(1)} = (2, 2, 0, 0)$, $x^{(2)} = (3, 0, 6, 0)$, $x^{(5)} = (0, 3, 0, 3)$ and $x^{(6)} = (0, 0, 12, 6)$. Two other solutions are **infeasible**[12].

Example 2
Solve the system of inequalities:

$$\begin{cases} x_1 + 3x_2 - x_3 \geq 3, \\ 2x_1 + 7x_2 + x_3 \leq 5, \\ x_1 + 4x_2 + 2x_3 \leq 4. \end{cases}$$

Solution
After introducing the slack variables, the system takes the form:

$$\begin{cases} x_1 + 3x_2 - x_3 - x_4 = 3, \\ 2x_1 + 7x_2 + x_3 + x_5 = 5, \\ x_1 + 4x_2 + 2x_3 + x_6 = 4. \end{cases}$$

[12] Sometimes, in order to simplify the considerations, we call the basic feasible solutions simply "basic solutions" and we skip the infeasible ones.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that the coefficient by the slack variable x_4 equals to –1, as the first inequality has the form "≥". The augmented matrix and the reduction process have been presented below:

$$\left[\begin{array}{ccc|ccc|c} 1 & 3 & -1 & -1 & 0 & 0 & 3 \\ 2 & 7 & 1 & 0 & 1 & 0 & 5 \\ 1 & 4 & 2 & 0 & 0 & 1 & 4 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \rightarrow \\ w_3' = w_3 - w_1 \end{array}$$

$$\rightarrow \left[\begin{array}{ccc|ccc|c} 1 & 3 & -1 & -1 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 & 0 & 1 & 1 \end{array}\right] \begin{array}{l} w_1' = w_1 - 3w_2 \\ w_2' = w_2 \\ w_3' = w_3 - w_2 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{ccc|ccc|c} 1 & 0 & -10 & -7 & -3 & 0 & 6 \\ 0 & 1 & 3 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 1 & 2 \end{array}\right].$$

The first segment of the matrix reached its reduced form, but there is a zeroed row in it (the third one). This means that we need to check the consistency of the system. For this purpose, we write the matrix consisting of the remaining elements of the zeroed rows (i.e. the elements lying in the segments two and three). In our case, this means that we are interested in the matrix consisting only of the last two segments of the third row of the reduced matrix:

$$\left[\begin{array}{ccc|c} -1 & -1 & 1 & 2 \end{array}\right]$$

(note that its rank is 1). The next step is writing the system of equations corresponding to this matrix. In our case it consists of one equation only: $-x_4 - x_5 + x_6 = 2$. Now we are trying to find at least one basic feasible solution of this smaller system (here: equation). The considered one-equation system has three basic solutions: $x^{(1)} = (-2, 0, 0)$ corresponding to the base $B^{(1)} = \{x_4\}$, $x^{(2)} = (0, -2, 0)$ corresponding to the base $B^{(2)} = \{x_5\}$ and $x^{(3)} = (0, 0, 2)$ corresponding to the base $B^{(3)} = \{x_6\}$. As we can see, the third one ($x_4 = 0$, $x_5 = 0$, $x_6 = 2$) consists of nonnegative variables and thus is feasible. It means that the initial system of inequalities is consistent.

We find the general solution in those rows, in which there was at least one nonzero element in the first segment. In this example these are all the rows but the third one, thus the first one and the second. The reduced system has the form:

$$\begin{cases} x_1 - 10x_3 - 7x_4 - 3x_5 = 6, \\ x_2 + 3x_3 + 2x_4 + x_5 = -1. \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The identity matrix is formed by the columns corresponding to the variables x_1 and x_2, so the other ones are parameters. If we denote x_3, x_4 and x_5 by a, b_1 and b_2, respectively, then we can write the general solution in the following form: $x_1 = 6 + 10a + 7b_1 + 3b_2$, $x_2 = -1 - 3a - 2b_1 - b_2$, $x_3 = a$. We have to remember about the assumptions: $a \in R$, $b_1 \in R_+ \cup \{0\}$, $b_2 \in R_+ \cup \{0\}$. The last two assumptions originate in the fact that $x_4 \geq 0$ and $x_5 \geq 0$. We still have not taken into account the last one: $x_6 \geq 0$. In order to do that we will use again the equation: $-x_4 - x_5 + x_6 = 2$. We can find the formula for the missing variable: $x_6 = 2 + x_4 + x_5 = 2 + b_1 + b_2$ and write the last assumption in the form $2 + b_1 + b_2 \geq 0$. Finally the general solution is as follows:

$$\begin{cases} x_1 = 6 + 10a + 7b_1 + 3b_2, \\ x_2 = -1 - 3a - 2b_1 - b_2, \\ x_3 = a, \end{cases}$$
$$a \in R, \quad b_1, b_2 \geq 0, \quad 2 + b_1 + b_2 \geq 0.$$

All the basic solutions, feasible (corresponding to the bases $B^{(4)}$, $B^{(7)}$, $B^{(13)}$ and $B^{(19)}$) and infeasible, have been presented in Table 5.1. Observe that in one case (base $B^{(1)}$) it was impossible to find the solution, as the system resulting from substituting 0 to the non-basic variables was inconsistent. Note also that degeneration that occurred in some cases.

Table 5.1. Basic solutions

Base	Solution	Base	Solution
$B^{(1)} = \{x_1, x_2, x_3\}$	contradiction	$B^{(11)} = \{x_2, x_3, x_4\}$	(0, 9, –2, –2, 0, 0)
$B^{(2)} = \{x_1, x_2, x_4\}$	(–8, 3, 0, –2, 0, 0)	$B^{(12)} = \{x_2, x_3, x_5\}$	(0, 1, 0, 0, –2, 0)
$B^{(3)} = \{x_1, x_2, x_5\}$	(0, 1, 0, 0, –2, 0)	$B^{(13)} = \{x_2, x_3, x_6\}$	(0, 4/5, –3/5, 0, 0, 2)
$B^{(4)} = \{x_1, x_2, x_6\}$	(6, –1, 0, 0, 0, 2)	$B^{(14)} = \{x_2, x_4, x_5\}$	(0, 1, 0, 0, –2, 0)
$B^{(5)} = \{x_1, x_3, x_4\}$	(2, 0, 1, –2, 0, 0)	$B^{(15)} = \{x_2, x_4, x_6\}$	(0, 5/7, 0, –6/7, 0, 8/7)
$B^{(6)} = \{x_1, x_3, x_5\}$	(10/3, 0, 1/3, 0, –2, 0)	$B^{(16)} = \{x_2, x_5, x_6\}$	(0, 1, 0, 0, –2, 0)
$B^{(7)} = \{x_1, x_3, x_6\}$	(8/3, 0, –1/3, 0, 0, 2)	$B^{(17)} = \{x_3, x_4, x_5\}$	(0, 0, 2, –5, 3, 0)
$B^{(8)} = \{x_1, x_4, x_5\}$	(4, 0, 0, 1, –3, 0)	$B^{(18)} = \{x_3, x_4, x_6\}$	(0, 0, 5, –8, 0, –6)
$B^{(9)} = \{x_1, x_4, x_6\}$	(5/2, 0, 0, –1/2, 0, 3/2)	$B^{(19)} = \{x_3, x_5, x_6\}$	(0, 0, –3, 0, 8, 10)
$B^{(10)} = \{x_1, x_5, x_6\}$	(3, 0, 0, 0, –1, 1)	$B^{(20)} = \{x_4, x_5, x_6\}$	(0, 0, 0, –3, 5, 4)

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
Solve the system of inequalities:

$$\begin{cases} x_1 + 2x_2 \le 4, \\ 2x_1 + 5x_2 \le 5, \\ x_1 + 5x_2 \ge 4, \\ x_1 + x_2 \ge 4. \end{cases}$$

Solution
After introducing the slack variables, the system looks as follows:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4, \\ 2x_1 + 5x_2 + x_4 = 5, \\ x_1 + 5x_2 - x_5 = 4, \\ x_1 + x_2 - x_6 = 4. \end{cases}$$

The augmented matrix and the reduction process are presented below:

$$\left[\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 0 & 0 & 4 \\ 2 & 5 & 0 & 1 & 0 & 0 & 5 \\ 1 & 5 & 0 & 0 & -1 & 0 & 4 \\ 1 & 1 & 0 & 0 & 0 & -1 & 4 \end{array}\right] \begin{array}{l} w_1' = w_1 \\ w_2' = w_2 - 2w_1 \\ w_3' = w_3 - w_1 \\ w_4' = w_4 - w_1 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{cccccc|c} 1 & 2 & 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & -2 & 1 & 0 & 0 & -3 \\ 0 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & -1 & 0 \end{array}\right] \begin{array}{l} w_1' = w_1 - 2w_2 \\ w_2' = w_2 \\ w_3' = w_3 - 3w_2 \\ w_4' = w_4 + w_2 \end{array} \rightarrow$$

$$\rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 5 & -2 & 0 & 0 & 10 \\ 0 & 1 & -2 & 1 & 0 & 0 & -3 \\ 0 & 0 & 5 & -3 & -1 & 0 & 9 \\ 0 & 0 & -3 & 1 & 0 & -1 & -3 \end{array}\right].$$

As in the previous example, there are zeroed rows in the first segment (the third and the fourth). Let us write the matrix of the smaller system:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\left[\begin{array}{cccc|c} 5 & -3 & -1 & 0 & 9 \\ -3 & 1 & 0 & -1 & -3 \end{array}\right].$$

The system itself has the following form:

$$\begin{cases} 5x_3 - 3x_4 - x_5 = 9, \\ -3x_3 + x_4 - x_6 = -3. \end{cases}$$

Al the basic solutions of this system have been presented in Table 5.2.

Table 5.2. Basic solutions

Base	Solution	Base	Solution
$B^{(1)} = \{x_3, x_4\}$	$x^{(1)} = (0, -3, 0, 0)$	$B^{(4)} = \{x_4, x_5\}$	$x^{(4)} = (0, -3, 0, 0)$
$B^{(2)} = \{x_3, x_5\}$	$x^{(2)} = (1, 0, -9/4, 0)$	$B^{(5)} = \{x_4, x_6\}$	$x^{(5)} = (0, -3, 0, 0)$
$B^{(3)} = \{x_3, x_6\}$	$x^{(3)} = (9/5, 0, 0, -12/5)$	$B^{(6)} = \{x_5, x_6\}$	$x^{(6)} = (0, 0, -9, 3)$

None of the solutions is feasible – at least one variable is negative in each of them. This means that the initial system of inequalities is inconsistent.

EXERCISES

1. Solve the following systems of linear inequalities:

a) $\begin{cases} 2x_1 + x_2 + 3x_3 \le 20, \\ 5x_1 + 3x_2 + 8x_3 \le 10; \end{cases}$

b) $\begin{cases} x_1 - 3x_2 + 7x_3 \le 15, \\ 3x_1 - 8x_2 + 20x_3 \le 15; \end{cases}$

c) $\begin{cases} x_1 + 3x_2 \le 12, \\ 2x_1 + 7x_2 \le 18, \\ 5x_1 + 17x_2 \ge 10; \end{cases}$

d) $\begin{cases} 3x_1 + 4x_2 \le 24, \\ 9x_1 + 14x_2 \ge 28, \\ 6x_1 + 9x_2 \ge 10; \end{cases}$

e) $\begin{cases} x_1 - 3x_2 + 7x_3 \le 21, \\ 3x_1 - 9x_2 + 21x_3 \ge 30; \end{cases}$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

f) $\begin{cases} 3x_1 + 2x_2 - 7x_3 \le 40, \\ -6x_1 - 4x_2 + 14x_3 \le 100; \end{cases}$

g) $\begin{cases} 2x_1 - 5x_2 + 7x_3 \le 14, \\ 5x_1 - 7x_2 + 11x_3 \ge 16, \\ 3x_1 - 13x_2 + 17x_3 \ge 50; \end{cases}$

h) $\begin{cases} 7x_1 + 2x_2 \le 20, \\ x_1 - 4x_2 \ge 6, \\ 3x_1 + 3x_2 \ge 7, \\ 2x_1 + 7x_2 \ge 5. \end{cases}$

SOLUTIONS

1. Sample general solutions and basic solutions: feasible (b.f.s.) and infeasible (b.i.s.):
   a) $x_1 = 50 - a - 3b_1 + b_2$, $x_2 = -80 - a + 5b_1 - 2b_2$, $x_3 = a$, where $a \in R$, $b_1, b_2 \ge 0$;
   b.f.s.: (50, –80, 0, 0, 0), (130, 0, –80, 0, 0), (2, 0, 0, 16, 0), (0, –130, 50, 0, 0), (0, 10/3, 0, 50/3, 0), (0, 0, 5/4, 65/4, 0), (0, 0, 0, 20, 10);
   b.i.s.: (10, 0, 0, 0, –40), (0, 20, 0, 0, –50), (0, 0, 20/3, 0, –50/3).
   b) $x_1 = -75 - 4a + 8b_1 - 3b_2$, $x_2 = -30 + a + 3b_1 - b_2$, $x_3 = a$, where $a \in R$, $b_1, b_2 \ge 0$;
   b.f.s.: (–75, –30, 0, 0, 0), (–195, 0, 30, 0, 0), (5, 0, 0, 10, 0), (0, –195/4, –75/4, 0, 0), (0, 15/8, 0, 75/8, 0), (0, 0, 3/4, 39/4, 0), (0, 0, 0, 15, 15);
   b.i.s.: (15, 0, 0, 0, –30), (0, –5, 0, 0, –25), (0, 0, 15/7, 0, –195/7).
   c) $x_1 = 30 - 7b_1 + 3b_2$, $x_2 = -6 + 2b_1 - b_2$, where $b_1, b_2 \ge 0$, $b_1 + 2b_2 \le 38$;
   b.f.s.: (–236, 70, 38, 0, 0), (–27, –25, 0, 19, 0), (30, –6, 0, 0, 38), (2, 0, 10, 14, 0), (9, 0, 3, 0, 35), (0, 10/17, 174/17, 236/17, 0), (0, 18/7, 30/7, 0, 236/7), (0, 0, 12, 18, 10);
   b.i.s.: (12, 0, 0, –6, 50), (0, 4, 0, –10, 58).
   d) $x_1 = 112/3 - 7/3b_1 - 2/3b_2$, $x_2 = -22 + 3/2b_1 + 1/2b_2$, where $b_1, b_2 \ge 0$, $1/2b_1 - 1/2b_2 \le 16$;
   b.f.s.: (–112/3, 26, 32, 0, 0), (112/3, –22, 0, 0, 16), (28/9, 0, 44/3, 0, 26/3), (8, 0, 0, 44, 38), (0, 2, 16, 0, 8), (0, 6, 0, 56, 44);
   b.i.s.: (176/3, –38, 0, –32, 0), (5/3, 0, 19, –13, 0), (0, 10/9, 176/9, –112/9, 0), (0, 0, 24, –28, –10).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

e) $x_1 = 21 + 3a_1 - 7a_2 - b$, $x_2 = a_1$, $x_3 = a_2$, where $a_1, a_2 \in R$, $b \geq 0$, $3b \leq 33$;
b.f.s.: (10, 0, 0, 11, 0), (21, 0, 0, 0, 33),
(0, –10/3, 0, 11, 0), (0, –7, 0, 0, 33), (0, 0, 10/7, 11, 0);
b.i.s.: (0, 0, 3, 0, –33), (0, 0, 0, 21, –30),
in the case of bases $\{x_1, x_2\}$, $\{x_1, x_3\}$ and $\{x_2, x_3\}$: contradiction;
f) the system is inconsistent;
g) the system is inconsistent;
h) the system is inconsistent.

5.3. FOUNDATIONS OF LINEAR PROGRAMMING

THEORY IN A NUTSHELL

Making economic decisions usually involves the optimization of certain values (for example, profit maximization or cost minimization). This optimized quantity is called the **objective function** (or simply the **objective**). The decision maker has usually limited resources (time, materials, labor, etc.). These restrictions can be usually written in the form of inequalities, called **constraints**. **Feasible solution** is a solution that satisfies all constraints. The **optimal solution** is such a feasible solution, for which the value of the objective function is optimal (maximum or minimum). Sometimes the decision problem may not have optimal solutions (for example, if the system of constraints is inconsistent). **Linear programming** problem is a mathematical model of the decision problem with the property that the objective function is linear, and all constraints are linear inequalities or linear equations. The system of linear inequalities has usually infinitely many feasible solutions, so it is not possible to check all of them directly. Fortunately, the linear programming problems have a useful property.

Theorem 5.1. Optimal solutions of a linear programming problem
If a linear programming problem has an optimal solution, then at least one of the optima is a basic feasible solution of the system of constraints.

It follows that in order to find an optimal solution it is enough to find all its basic feasible solutions, find the value of the objective at each of them and choose the one where the objective value is optimal (maximum or minimum)[13].

[13] It can happen sometimes that a linear programming problem has infinitely many optimal solutions or that it has no optimal solution, although it has feasible solutions. We will not be interested, however, in such kind of problems.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to find the solution of a linear programming problem using **WolframAlpha®**, you can use the syntax "maximize/minimize {function} over {constraints separated with commas}". In particular, in order to solve the problem from Example 1 below, you can type:

"maximize {4x1+x2} over {3x1+x2<=12, 2x1+6x2<=24, x1>=0, x2>=0}".

EXAMPLES

Example 1

Solve the following linear programming problem:

$$(0)\ f(x_1, x_2) = 4x_1 + x_2 \to \max,$$
$$(1)\ 3x_1 + x_2 \le 12,$$
$$(2)\ 2x_1 + 6x_2 \le 24,$$
$$(3)\ x_1, x_2 \ge 0.$$

Solution

Note that the constraints and the objective have been numbered. It is not necessary, but recommended, as it will make their identification easier. The objective has number (0), the constraints – the next consecutive numbers. The last condition is usually the **nonnegativity constraint**. Its presence is rooted in the nature of problems that you will solve – for example, you cannot produce a negative number of products.

We begin the solution process with introducing the slack variables. Recall that we introduce one slack variable to each inequality. We do not introduce slack variables to the objective function. We also remember that the slack variables must be nonnegative. Taking this into account, we obtain the following form of the problem:

$$(0)\ f(x_1, x_2) = 4x_1 + x_2 \to \max,$$
$$(1)\ 3x_1 + x_2 + x_3 = 12,$$
$$(2)\ 2x_1 + 6x_2 + x_4 = 24,$$
$$(3)\ x_1, x_2, x_3, x_4 \ge 0.$$

The augmented matrix is as follows:

$$\left[\begin{array}{cc|cc|c} 3 & 1 & 1 & 0 & 12 \\ 2 & 6 & 0 & 1 & 24 \end{array}\right].$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that even without the elimination, we can conclude that the rank of the matrix is equal to 2. We need not determine the general solution and only the basic solutions, thus performing elementary operations is not necessary at all. Since the rank of the matrix is equal to 2, we need to determine the solutions corresponding to all pairs of variables. The results are shown in Table 5.3. The values of the objective are given in the last column (OBJ) corresponding to the solutions. For example, for the solution $x^{(1)}$ we have: $f(3, 3) = 4 \cdot 3 + 3 = 15$, for the solution $x^{(2)}$ we do not evaluate the objective as this solution is infeasible ($x_3 = -24 < 0$) and we do not take it into account. As we can see, the optimal (maximum) value of the objective corresponds to the solution $x^{(3)}$. Thus the optimal values of the variables are $x_1 = 4$ and $x_2 = 0$, and the corresponding maximum of the objective is 16. This is written in short as follows: $f_{\max}(4, 0) = 16$.

Table 5.3. Basic solutions and the objective (OBJ)

Base	Solution	OBJ	Base	Solution	OBJ
$B^{(1)} = \{x_1, x_2\}$	$x^{(1)} = (3, 3, 0, 0)$	15	$B^{(4)} = \{x_2, x_3\}$	$x^{(4)} = (0, 4, 8, 0)$	4
$B^{(2)} = \{x_1, x_3\}$	$x^{(2)} = (12, 0, -4, 0)$	×	$B^{(5)} = \{x_2, x_4\}$	$x^{(5)} = (0, 12, 0, -48)$	×
$B^{(3)} = \{x_1, x_4\}$	$x^{(3)} = (4, 0, 0, 16)$	16	$B^{(6)} = \{x_3, x_4\}$	$x^{(6)} = (0, 0, 12, 24)$	0

Example 2

The company produces two products P_1 and P_2, a process that uses two resources: S_1 and S_2. Standards of consumption of the production (in kg / pc.), unit profits (PLN) and weekly resource limits (kg) are presented in Table 5.4. Determine the volume of production that optimizes the weekly profit of the company.

Table 5.4. Information on products and resources

Product / Resource	P_1	P_2	Limit
S_1	3	6	1500
S_2	4	3	1200
Profit	10	9	

Solution

We begin by writing the above problem as a linear programming problem. The decision to be taken concerns the production volume for each product. This means that the decision variables should represent the volume of production. For example: x_1 – the volume of production of the product P_1 expressed in pieces, x_2 – the production volume of the product P_2 expressed in pieces. The objective function will represent the profit of the company, just as we want to maximize this value. Because profit from producing one piece of product P_1 is 10 PLN, so

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

the total profit from the production of P_1 will be $10x_1$ (profit from production of one piece equal to 10 PLN times the number of units, denoted by x_1). Similarly, the profit of the manufacture of the product P_2 is $9x_2$. The total profit is therefore:

$f(x_1, x_2) = 10x_1 + 9x_2$.

Let us now turn to the resources. Consumption of S_1 per unit of product P_1 is 3 kg. Hence, the total consumption of this resource to the production of P_1 is equal to $3x_1$ (consumption per piece equal to 3 kg times the number of pieces, denoted by x_1). Similarly, we calculate the consumption of S_1 to produce P_2: $6x_2$. Finally the total consumption of S_1 is $3x_1 + 6x_2$. Consumption cannot exceed the limit which is equal to 1500 kg. This can be written in the form of inequality: $3x_1 + 6x_2 \le 1500$. Similarly we can write the constraint corresponding with the second resource: $4x_1 + 3x_2 \le 1200$. Remember also that the variables must be nonnegative (you cannot produce a negative number of products). Considering the above, the corresponding the linear programming problem is as follows:

$$
\begin{aligned}
&(0)\ f(x_1, x_2) = 10x_1 + 9x_2 \to \max,\\
&(1)\ 3x_1 + 6x_2 \le 1500,\\
&(2)\ 4x_1 + 3x_2 \le 1200,\\
&(3)\ x_1, x_2 \ge 0.
\end{aligned}
$$

After introducing the slack variables it takes the form:

$$
\begin{cases}
3x_1 + 6x_2 + x_3 = 1500,\\
4x_1 + 3x_2 + x_4 = 1200,
\end{cases}
$$

and the augmented matrix is:

$$
\left[\begin{array}{cc|cc|c}
3 & 6 & 1 & 0 & 1500\\
4 & 3 & 0 & 1 & 1200
\end{array}\right].
$$

The rank of the matrix is 2, so all the bases will consists of two elements. All the bases, the corresponding basic solutions and objective values (OBJ) have been presented in Table 5.5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Table 5.5. Basic solutions and the objective (OBJ)

Base	Solution	OBJ
$B^{(1)} = \{x_1, x_2\}$	$x^{(1)} = (180, 160, 0, 0)$	3240
$B^{(2)} = \{x_1, x_3\}$	$x^{(2)} = (300, 0, 600, 0)$	3000
$B^{(3)} = \{x_1, x_4\}$	$x^{(3)} = (500, 0, 0, -800)$	×
$B^{(4)} = \{x_2, x_3\}$	$x^{(4)} = (0, 400, -900, 0)$	×
$B^{(5)} = \{x_2, x_4\}$	$x^{(5)} = (0, 250, 0, 450)$	2250
$B^{(6)} = \{x_3, x_4\}$	$x^{(6)} = (0, 0, 1500, 1200)$	0

As we can see, the maximum value of the objective corresponds to $x^{(1)}$: $f_{\max}(180, 160) = 3240$. It follows that in order to achieve the maximum weekly profit, equal to 3240 PLN, the company should produce 180 pieces of P_1 and 160 pieces of P_2 weekly.

Example 3

The cat can be fed with dry food or canned food. The water and protein content in the two types of food (grams per 100 g), their prices (PLN per 100 g) and the required daily intake of both components (in grams) are given in Table 5.6. Determine the daily cat diet with minimal cost.

Table 5.6. Information about food and contents

Content \ Food	Canned	Dry	Required intake
Protein	10	40	30
Water	80	0	80
Price per 100 g	3	2	

Solution

This time the decision variables will be: x_1 – the amount of 100 g – servings of the canned food in the daily cat diet, x_2 – the amount of 100 g – servings of the dry food in the daily cat diet. The daily cost will be: $3x_1 + 2x_2$, the protein intake $10x_1 + 40x_2$, and the water intake – $80x_1$. The intake of both ingredients must be at least as high as the required amount, so both inequalities will take the form "≥". The cost of the food should be as small as possible. This means that the corresponding linear programming problem is:

$$
\begin{aligned}
&(0)\ f(x_1, x_2) = 3x_1 + 2x_2 \to \min,\\
&(1)\ 10x_1 + 40x_2 \geq 30,\\
&(2)\ 80x_1 \geq 80,\\
&(3)\ x_1, x_2 \geq 0.
\end{aligned}
$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

After introducing the slack variables it takes the form:

$$\begin{cases} 10x_1 + 40x_2 - x_3 = 30, \\ 80x_1 - x_4 = 80, \end{cases}$$

and the augmented matrix is (note that its rank equals to 2):

$$\left[\begin{array}{cc|cc|c} 10 & 40 & -1 & 0 & 30 \\ 80 & 0 & 0 & -1 & 80 \end{array}\right].$$

The basic solutions and corresponding values of the objective are presented in Table 5.7.

Table 5.7. Basic solutions and the objective (OBJ)

Base	Solution	OBJ
$B^{(1)} = \{x_1, x_2\}$	$x^{(1)} = (1, 1/2, 0, 0)$	4
$B^{(2)} = \{x_1, x_3\}$	$x^{(2)} = (1, 0, -20, 0)$	×
$B^{(3)} = \{x_1, x_4\}$	$x^{(3)} = (3, 0, 0, 160)$	9
$B^{(4)} = \{x_2, x_3\}$	contradiction	×
$B^{(5)} = \{x_2, x_4\}$	$x^{(5)} = (0, 3/4, 0, -80)$	×
$B^{(6)} = \{x_3, x_4\}$	$x^{(6)} = (0, 0, -30, -80)$	×

The smallest value of the objective corresponds to $x^{(1)}$: $f_{\min}(1, 1/2) = 4$. It means that in order to achieve the minimum cost of feeding the cat, equal to 4 PLN, it is necessary to give him every day 1 serving (100 g) of canned food and 1/2 serving (50 g) of dry food.

EXERCISES

1. Solve the linear programming problems:

a) (0) $f(x_1, x_2) = 5x_1 + 6x_2 \to \max$,
(1) $4x_1 + x_2 \le 20$,
(2) $2x_1 + 8x_2 \le 40$,
(3) $x_1, x_2 \ge 0$;

b) (0) $f(x_1, x_2) = 5x_1 + 6x_2 \to \min$,
(1) $4x_1 + x_2 \ge 20$,
(2) $2x_1 + 8x_2 \ge 40$,
(3) $x_1, x_2 \ge 0$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) (0) $f(x_1, x_2) = 3x_1 + 10x_2 \to \max$,
(1) $2x_1 + x_2 \le 24$,
(2) $3x_1 + 6x_2 \le 72$,
(3) $x_1, x_2 \ge 0$;

d) (0) $f(x_1, x_2) = 3x_1 + 10x_2 \to \min$,
(1) $2x_1 + x_2 \ge 24$,
(2) $3x_1 + 6x_2 \ge 72$,
(3) $x_1, x_2 \ge 0$.

2. The company produces two products using two resources. Data on unit profit (in thousand PLN), consumption of resources (in kg per piece), and monthly resource limits (kg) are presented in the table. Determine the production plan maximizing the profit of the company.

Product / Resource	P_1	P_2	Limit
S_1	5	10	150
S_2	20	10	300
Profit	4	3	

3. The company produces two products using two resources. Data on unit profit (in thousand PLN), consumption of resources (in kg per piece), and monthly resource limits (kg) are presented in the table. Determine the production plan maximizing the profit of the company.

Product / Resource	P_1	P_2	Limit
S_1	5	15	240
S_2	30	10	480
Profit	3	4	

4. The diet consists of two food products that comprise, among others, two nutrients. Data on prices of products (in PLN per 100 g), the content of the ingredients in the products (g per 100 g) and the weekly norm of nutrient intake (g) are presented in the table. Find a diet with minimal cost.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Product / Nutrient	P_1	P_2	Required amount
S_1	5	10	150
S_2	20	10	300
Price	4	3	

5. The diet consists of two food products that comprise, among others, two nutrients. Data on prices of products (in PLN per 100 g), the content of the ingredients in the products (g per 100 g) and the weekly norm of nutrient intake (g) are presented in the table. Find a diet with minimal cost.

Product / Nutrient	P_1	P_2	Required amount
S_1	5	15	240
S_2	30	10	480
Price	3	4	

SOLUTIONS

1. a) $f_{\max}(4, 4) = 44$; b) $f_{\min}(4, 4) = 44$; c) $f_{\max}(0, 12) = 120$; d) $f_{\min}(24, 0) = 72$.
2. x_1 – the volume of production of the product P_1 expressed in pieces, x_2 – the volume of production of the product P_2 expressed in pieces; the model:

   (0) $f(x_1, x_2) = 4x_1 + 3x_2 \to \max$,

   (1) $5x_1 + 10x_2 \le 150$,

   (2) $20x_1 + 10x_2 \le 300$,

   (3) $x_1, x_2 \ge 0$.

   Optimal volume of production: 10 pc. of P_1 and 10 pc. of P_2, maximum profit: 70 000 PLN.

3. x_1 – the volume of production of the product P_1 expressed in pieces, x_2 – the volume of production of the product P_2 expressed in pieces; the model:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

(0) $f(x_1, x_2) = 3x_1 + 4x_2 \to \max$,

(1) $5x_1 + 15x_2 \le 240$,

(2) $30x_1 + 10x_2 \le 480$,

(3) $x_1, x_2 \ge 0$.

Optimal volume of production: 12 pc. of P_1 and 12 pc. of P_2, maximum profit: 84 000 PLN.

4. x_1 – the amount of 100 g – servings of P_1, x_2 – the amount of 100 g – servings of P_2, the model:

(0) $f(x_1, x_2) = 4x_1 + 3x_2 \to \min$,

(1) $5x_1 + 10x_2 \ge 150$,

(2) $20x_1 + 10x_2 \ge 300$,

(3) $x_1, x_2 \ge 0$.

Optimal diet: 1 kg of P_1 and 1 kg of P_2, minimum cost: 70 PLN.

5. x_1 – the amount of 100 g – servings of P_1, x_2 – the amount of 100 g – servings of P_2, the model:

(0) $f(x_1, x_2) = 3x_1 + 4x_2 \to \min$,

(1) $5x_1 + 15x_2 \ge 240$,

(2) $30x_1 + 10x_2 \ge 480$,

(3) $x_1, x_2 \ge 0$.

Optimal diet: 1,2 kg of P_1 and 1,2 kg of P_2, minimum cost: 84 PLN.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 6
BASIC FUNCTIONS AND THEIR PROPERTIES

6.1. BASIC CONCEPTS

THEORY IN A NUTSHELL

In this section some basic concepts related to functions have been introduced. We will not use any formal definitions, and only intuitive descriptions to help you understand them. Examples and exercises have been neglected, because the study of various properties of functions will need some deeper knowledge and therefore it will be carried out in the following sections and chapters.

Let us recall[14], that **function is such a relation that each member of the domain is related to at most one member of the counterdomain**. For that reason instead of xfy we write simply $y = f(x)$ (because y is defined unambiguously).
A function is **increasing**, if its values y increase together with its arguments x, and **decreasing**, if y decrease when x increase. If the output values y of f are the same for all the input values x, then we say that f is a **constant function**. Sample plots of increasing, decreasing and constant functions are presented in Fig. 6.1.

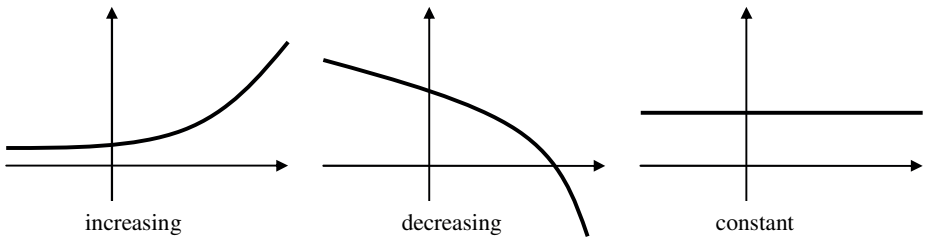


Fig. 6.1. Increasing, decreasing and constant function

Local extremum is the point, where the function takes the highest (**local maximum**) or lowest (**local minimum**) value within a given neighborhood. Note that a local maximum (minimum) is not necessarily the point where the function takes the **globally** highest (lowest) value. It is the point where the function takes the **locally** highest (lowest) value. The examples of local maximum and minimum are presented in Fig. 6.2.

[14] See Definition 3.3 in Chapter 3 "Relations".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

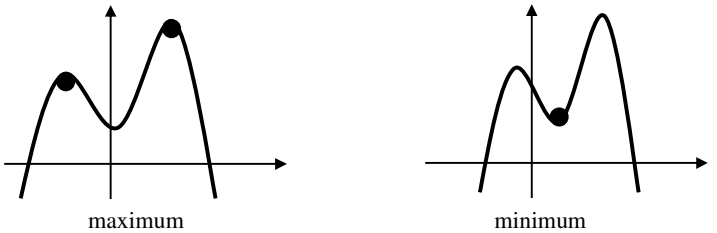


Fig. 6.2. Local maximum and local minimum

6.2. POLYNOMIALS AND RATIONAL FUNCTIONS

THEORY IN A NUTSHELL

Polynomial is a function in the form $W(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$. **The degree** of a polynomial is the highest exponent of any power of x (the polynomial in the above formula has degree n). In the general case (omitting polynomial of degree 0, which is a constant function) values of a polynomial function increase to ∞ or decrease to $-\infty$, when x increases or decreases boundlessly. A polynomial cannot have more zeros (i.e., different values of x for which the value of function is 0; sometimes called **roots**) than its degree. Usually polynomial function alternately increases and decreases, having no more than $n - 1$ extrema, where n is the degree. Sample plot of a polynomial has been presented in Fig. 6.3.

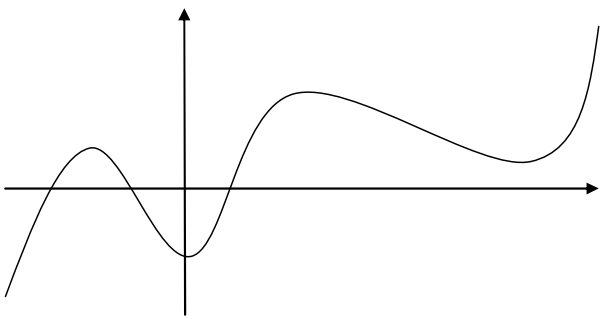


Fig. 6.3. Polynomial function

In the remainder of this section we will focus on polynomial and rational equations and inequalities.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Linear equation is an equation of the form $ax + b = 0$. It has exactly one solution provided that $a \neq 0$. This solution equals to $x = -b / a$.
Quadratic equation is an equation of the form $ax^2 + bx + c = 0$ (we assume that $a \neq 0$). Its **discriminant** you will derive from the formula $\Delta = b^2 - 4ac$. If $\Delta > 0$, then the equation has two distinct solutions: $x_1 = \frac{-b-\sqrt{\Delta}}{2a}$, $x_2 = \frac{-b+\sqrt{\Delta}}{2a}$ and can be written in an equivalent form $a(x - x_1)(x - x_2) = 0$. If $\Delta = 0$, then the equation has only one solution $x_0 = \frac{-b}{2a}$ and can be written in the form $a(x - x_0)^2 = 0$. If $\Delta < 0$, then the equation does not have solutions[15].
Polynomial equation is an equation of the form $W(x) = 0$, where $W(x)$ is a polynomial. We solve equations of this type in one of two ways: by **factorization**, or by using the following theorem:

Theorem 6.1. Rational Root Theorem
Assume that given is a polynomial $W(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ with integer coefficients. If the equation $W(x) = 0$ has rational roots of the form $x = p / q$ (where p and q are integers), then p is a divisor[16] of a_0 and q is a divisor of a_n.

Rational equation is an equation of the form $\frac{W(x)}{V(x)} = 0$. Its solutions are all the solutions of the equation $W(x) = 0$, excluding the solutions of the equation $V(x) = 0$ (the denominator cannot be equal to 0).
The linear, quadratic and polynomial inequalities of the form $W(x) > 0$, $W(x) < 0$, $W(x) \geq 0$ and $W(x) \leq 0$ are solved in this way, that first we solve the corresponding equation, then the roots of the respective polynomial (i.e. solutions of the equation) are marked on the real line. Then we check the sign of the coefficient of the highest power of x. We sketch the plot starting from the right side, starting from the top when the coefficient is positive, or from the bottom when it is negative. At each point corresponding to a solution, we check its multiplicity, and if it is even, the graph remains on the same side (above or below) the axis, otherwise we change the side. Finally, we read the intervals, for

[15] More precisely: it has no solutions in real numbers. If you have any experience with complex numbers, then you know that the quadratic equation has a solution even if the discriminant is negative. In this book, however, we will not deal with complex numbers, as they are rather not useful in economic issues that you will deal with in the future.
[16] The expression "m is a divisor of n" means that n is divisible by m, giving an integer as result. For example $18/(-3) = -6$ (an integer), so -3 is a divisor of 18. On the other hand $18/4 = 4{,}5$ (not an integer), thus 4 is not a divisor of 18.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

which the sketch is above or below the axis. Remember to take into account the ends of the intervals if the inequality is tight (i.e., "≤" or "≥").

Rational inequalities of the form $\frac{W(x)}{V(x)} > 0$, $\frac{W(x)}{V(x)} < 0$, $\frac{W(x)}{V(x)} \geq 0$ and $\frac{W(x)}{V(x)} \leq 0$ have the same solutions as the inequalities, $W(x)V(x) > 0$, $W(x)V(x) < 0$, $W(x)V(x) \geq 0$ and $W(x)V(x) \leq 0$, respectively (we use here the fact that the sign of the quotient of two numbers is the same as the sign of their product). It is only necessary to exclude all the solutions of the equation $V(x) = 0$ (the denominator cannot be equal to 0).

The absolute value of a number is the nonnegative value of this number, without regard to the sign. More formally, the absolute value of x is defined as:

$$|x| = \begin{cases} x, & \text{when } x \geq 0, \\ -x, & \text{when } x < 0. \end{cases}$$

You can incorporate 0 to the second part of the formula, if necessary. Equations and inequalities with absolute value solved by considering two cases: when the expression under the absolute value is negative (nonpositive) and positive (nonnegative). It is also worth remembering the following rules:

$$|x| = a \quad \Leftrightarrow \quad x \in \{-a, a\},$$

$$|x| \leq a \quad \Leftrightarrow \quad x \in [-a, a],$$

$$|x| \geq a \quad \Leftrightarrow \quad x \in (-\infty, -a] \cup [a, \infty).$$

If there are sharp inequalities instead of tight on the left hand side, then the intervals on the right hand side should be open.

If you want to solve any type of inequality or equation using **WolframAlpha®**, you can just type it. For example, if you want to solve the inequality from Example 4: $x^5 - 5x^4 - 4x^3 + 20x^2 > 0$, then simply type:

"x^5-5x^4-4x^3+20x^2>0".

If you want to find the domain of a function, then you should use the word "domain" and then give the formula. For example, to find the domain of $f(x) = \sqrt{2x - x^2}$, type:

"domain f(x)=sqrt(2x-x^2)"

(**sqrt** is notation for the **square root**).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1
Solve the equation: $2x + 7 = 12$.

Solution
We write the equation in its standard form: $2x - 5 = 0$, and then we use the formula: $x = -(-5)/2 = 5/2$.

Example 2
Solve the inequality: $2x + 7 \leq 12$.

Solution
From the previous example we know that the corresponding equation has solution $x = 5/2$. As the coefficient by x is positive, the solution is $x \leq 5/2$ (we do not change the direction; if the coefficient by x were negative, it would be necessary to change the direction).

Example 3
Solve the inequality: $2x^2 + 5x + 2 \leq 0$.

Solution
First we will solve the corresponding equation: $2x^2 + 5x + 2 = 0$. Its discriminant is equal to $\Delta = 5^2 - 4 \cdot 2 \cdot 2 = 9$. Thus we know that the equation has two solutions $x_1 = \frac{-5-\sqrt{9}}{2 \cdot 2} = -2$ and $x_2 = \frac{-5+\sqrt{9}}{2 \cdot 2} = -\frac{1}{2}$. The function $f(x) = 2x^2 + 5x + 2$ on the left side of the equation (and inequality) can be thus written in the form $f(x) = 2(x + 2)(x + 1/2)$. Observe that the exponent of both powers is 1. In order to solve the respective inequality, we will sketch the plot of the function $f(x) = 2x^2 + 5x + 2$. As the coefficient by the highest power of x (i.e., by x^2) is positive (equal to 2), we begin drawing from the top right, and then cross the axis twice (powers of both solutions have odd exponents). The sketch is presented in Fig. 6.4.

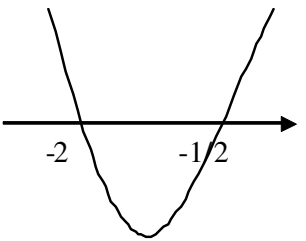


Fig. 6.4. Sketch of the function

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

As we can see, the function takes values lower than or equal to 0 when x is between –2 and –1/2, so the solution of the inequality has the form of the interval [–2, –1/2] (the inequality is tight, so the interval is closed).

Example 4
Solve the inequality: $x^5 - 5x^4 - 4x^3 + 20x^2 > 0$.

Solution
We will begin again by solving the corresponding equation. Observe that all the left hand side elements are divisible by x^2, so we factor x^2 out to the front of the parentheses: $x^5 - 5x^4 - 4x^3 + 20x^2 = x^2(x^3 - 5x^2 - 4x + 20)$. The two first elements in the bracket are now divisible by x^2, and two latter ones by 4: $x^2(x^3 - 5x^2 - 4x + 20) = x^2[x^2(x - 5) - 4(x - 5)]$. As we can see, it is also possible to factor out the expression $(x - 5)$. The remaining polynomial (i.e., $x^2 - 4$), can be factorized using one of the short multiplication formulae: $(x - 2)(x + 2) = x^2 - 2^2$. Finally, the left hand side takes the form: $f(x) = x^2(x - 5)(x - 2)(x + 2)$. It follows that the solutions are 0, 5, 2 and –2, where 0 is a double root, thus of an even multiplicity – the expression x, i.e., $(x - 0)$, is squared. The other solutions are simple roots. The coefficient by the highest power of x (i.e., x^5) is positive (equal to 1), so we start the sketch from the top right and then cross the axis at all the roots except 0 (there is a reflection at 0), as it has been shown in Fig. 6.5.

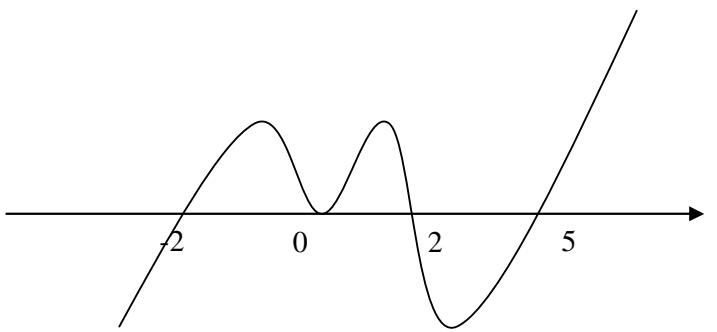


Fig. 6.5. Sketch of the function

The solution is sum of the intervals for which the plot lies over the axis, i.e. $(-2, 0) \cup (0, 2) \cup (5, \infty)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 5

Solve the inequality: $\frac{6x^3 - 11x^2 + 6x - 1}{(x-1)} \le 0$.

Solution

We begin by transforming quotient into product. This way we obtain inequality: $(6x^3 - 11x^2 + 6x - 1)(x - 1) \le 0$. The solutions of both inequalities are almost the same, it is only necessary to exclude the zeros of the denominator: $x - 1 \ne 0$, i.e. $x \ne 1$. Factoring out any expression is rather difficult this time, so we will use the Rational Root Theorem. The divisors of a_0 (i.e., 1) are –1 and 1, and the divisors of the coefficient by the highest power of x (i.e., 6) are –1, 1, –2, 2, –3 and 3. Let us try to find a quotient that would be a root of the polynomial in brackets. Let us begin with the first pair: $(-1)/(-1) = 1$. We substityute 1 to the expression in brackets and we obtain: $6 \cdot 1^3 - 11 \cdot 1^2 + 6 \cdot 1 - 1 = 0$, so 1 is the root that we were looking for (usually you will not find a solution so quickly). It follows that the polynomial can be written in the form $(x - 1) \cdot W(x)$. In order to find $W(x)$, we divide the expression in brackets by $(x - 1)$. For this purpose we divide the monomial wit the highest power of x by x: $6x^3/x = 6x^2$. But $6x^2(x - 1) = 6x^3 - 6x^2$, so we still have to divide $(6x^3 - 11x^2 + 6x - 1) - (6x^3 - 6x^2) = -5x^2 + 6x - 1$. We divide $-5x^2$ by x, and obtain $-5x$. But $-5x(x - 1) = -5x^2 + 5x$, so we still have to divide $(-5x^2 + 6x - 1) - (-5x^2 + 5x) = x - 1$, which in turn divided by $(x - 1)$ gives 1 as the result. Finally, the result of division is $6x^2 - 5x + 1$. It is a polynomial of degree 2 with roots $x_1 = 1/2$ and $x_2 = 1/3$, so finally the expression in the first brackets can be written as $6(x - 1)(x - 1/2)(x - 1/3)$, and the inequality takes finally the form $6(x - 1)^2(x - 1/2)(x - 1/3) \le 0$. The sketch of the corresponding function has been presented in Fig. 6.6.

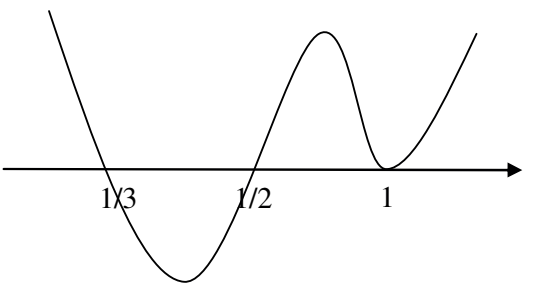


Fig. 6.6. Sketch of the function

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution of the last inequality has the form $[1/3, 1/2] \cup [1, \infty)$. We must remember, however, to exclude the number 1, so the solution of the initial inequality is $[1/3, 1/2] \cup (1, \infty)$.

Example 6
Solve the inequality: $|2x - 3| < 5$.

Solution
It follows from the properties of the absolute value that given inequality is satisfied if and only if two following inequalities are satisfied simultaneously: $2x - 3 < 5$ and $2x - 3 > -5$. The solution of the first one is $(-\infty, 4)$, and the solution of the other one $(-1, \infty)$, so the solution of the initial inequality is their intersection, i.e. $(-1, 4)$.

Example 7

Find the domain of the function $f(x) = \dfrac{x^2 - 11}{x + 23}$.

Solution
The domain of rational function is the set of real numbers R, excluding the zeros of the denominator. Thus $D(f) = R \setminus \{-23\}$.

Example 8
Find the domain of the function $f(x) = \sqrt{100 - x^2}$.

Solution
The square-rooted expression cannot be negative, so the following inequality must be satisfied: $100 - x^2 \geq 0$. After solving it we conclude that $D(f) = [-10, 10]$.

EXERCISES

1. Solve the inequalities:
   a) $2x^2 + 5x - 3 \geq 0$;
   b) $6x^3 - 13x^2 + x + 2 \leq 0$;
   c) $x^3 + 3x^2 - 9x - 27 < 0$;
   d) $2x^3 + 5x^2 - 32x - 80 > 0$;
   e) $(3x^3 + 5x^2 - 27x - 45) / (x + 4) \leq 0$;
   f) $(x^4 - 6x^3 + 13x^2 - 12x + 4) / (x + 3) \leq 0$;
   g) $|3x - 6| > 12$;
   h) $|2x - 7| \leq 9$;
   i) $|x^2 - 3x + 1| < 1$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Find the domains of the functions:
   a) $f(x) = (x^2 - 2x + 1) / (x - 3)$;
   b) $f(x) = (x - 3) / (x^2 - 4x + 3)$;
   c) $f(x) = \sqrt{6x^2 - x - 1}$;
   d) $f(x) = \sqrt{2x - x^2}$;
   e) $f(x) = (x-5)/\sqrt{2x - x^2}$.

SOLUTIONS

1. a) $x \in (-\infty, -3] \cup [1/2, \infty)$; b) $x \in (-\infty, -1/3] \cup [1/2, 2]$;
   c) $x \in (-\infty, -3) \cup (-3, 3)$; d) $x \in (-4, -5/2) \cup (4, \infty)$;
   e) $x \in (-4, -3] \cup [-5/3, 3]$; f) $x \in (-\infty, -3) \cup \{1, 2\}$;
   g) $x \in (-\infty, -2) \cup (6, \infty)$; h) $x \in [-1, 4]$;
   i) $x \in (0, 1) \cup (2, 3)$.

2. a) $D(f) = R \setminus \{3\}$; b) $D(f) = R \setminus \{1, 3\}$;
   c) $D(f) = (-\infty, -1/3] \cup [1/2, \infty)$; d) $D(f) = [0, 2]$;
   e) $D(f) = (0, 2)$.

6.3. EXPONENTIAL FUNCTION

THEORY IN A NUTSHELL

Exponential function is a function of the form $f(x) = a^x$, where a is parameter. We assume that $a > 0$. The behavior of the function depends on the value of a. If $a \in (0, 1)$, then the function is decreasing – when x tends to $-\infty$, its value increases to ∞, and when x increases to ∞, its value tends to 0 (but it does not reach 0). We are in an opposite situation, when $a > 1$. Then the function increases: when x tends to $-\infty$, it tends to 0 (but does not reach 0), and when x increases to ∞, it increases to ∞. The exponential function is constant when $a = 1$ (its value is then equal to 1).
No matter what is the value of a, the value of the function in 0 is always equal to 1: $f(0) = a^0 = 1$. Remember also that the exponential functions **takes only positive values**. Sample behavior of exponential function depending on the value of a has been illustrated in Fig. 6.7.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

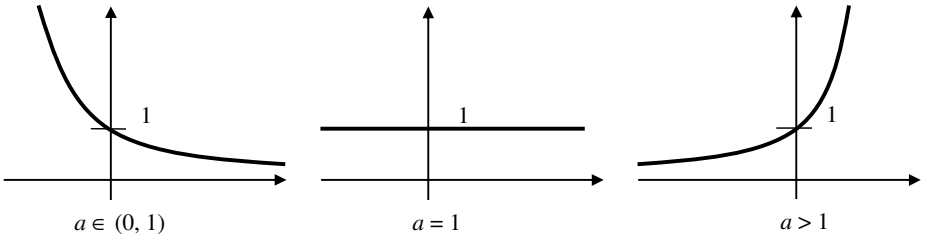


Fig. 6.7. Exponential function

Basic properties of the exponential function:

1. Operations on powers having the same exponent: $a^x b^x = (ab)^x$, $a^x / b^x = (a/b)^x$.
2. Operations on powers having the same base: $a^x a^y = a^{x+y}$, $a^x / a^y = a^{x-y}$.
3. Raising a power to a power: $(a^x)^y = a^{xy}$.

To solve an **exponential equation**, you must transform both sides to the form of powers having the same base, and then compare the exponents. In the case of an **exponential inequality** we proceed similarly, bearing in mind the fact that it is necessary to reverse the inequality, when $a < 1$ (note that the function decreases in this case).

From the point of view of applications, in the remainder of the book we will be particularly interested in the exponential function with parameter $a = e$ ($e = 2.71828\ldots$ is a mathematical constant called the **Euler's number** or the **Napier's constant**). The function $f(x) = e^x$ is often denoted by $f(x) = \exp(x)$.

In order to perform the operations with the exponential expressions using **WolframAlpha®**, you will need the symbol of power "^" and the function "exp". For example, if you want to simplify the expression $2^x / 4^x$, then you should use the code:

"simplify 2^x / 4^x".

If you want to solve the inequality

$e^{2x} - 1 > 0$,

then you could type:

"solve exp(2x)-1>0"

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

or

“solve e^(2x)-1>0”.

EXAMPLES

Example 1
Simplify the expression: $4^x \cdot 3^x / 6^x$.

Solution
We are dealing with the operations on the powers having the same exponent, so we will use the property 1:

$4^x \cdot 3^x / 6^x = (4 \cdot 3 / 6)^x = 2^x$.

Example 2
Simplify the expression: $(1/5)^x \cdot (1/5)^{3x+2} / (1/5)^{x-8}$.

Solution
This time we are dealing with the operations on the powers having the same base, so we will use the property 2:

$(1/5)^x \cdot (1/5)^{3x+2} / (1/5)^{x-8} = (1/5)^{x+(3x+2)-(x-8)} = (1/5)^{3x+10}$.

Example 3
Simplify the expression: $3^x \cdot 9^{2x+3}$.

Solution
As $9 = 3^2$, we can use the formula for raising a power to a power (property 3): $3^x \cdot 9^{2x+3} = 3^x \cdot (3^2)^{2x+3} = 3^x \cdot 3^{2 \cdot (2x+3)} = 3^x \cdot 3^{4x+6}$. We use also the property 2 (as we multiply the powers having the same base): $3^x \cdot 3^{4x+6} = 3^{x+(4x+6)} = 3^{5x+6}$.

Example 4
Solve the equation: $(1/3)^{x+5} = (1/9)^{2x-2}$.

Solution
We must begin by transforming both sides to the power of the same base. $1/9 = (1/3)^2$, so we will express both sides as the powers of base 1/3:

$(1/3)^{x+5} = [(1/3)^2]^{2x-2}$.

After applying the formula for raising a power to a power we obtain:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$(1/3)^{x+5} = (1/3)^{4x-4}$.

As both powers have the same base, we can compare the exponents :

$x + 5 = 4x - 4$.

It implies that $x = 3$.

Example 5
Solve the inequality: $(1/8)^{2x-1} < (1/2)^{x+2}$.

Solution
We bagin, as in the previous example, with expressing both sides as the powers of the same base. We use the identity $1/8 = (1/2)^3$:

$(1/2)^{6x-3} < (1/2)^{x+2}$.

Now we can compare the exponents. Note that we change the direction of the inequality because the base is lower than 1:

$6x - 3 > x + 2$.

Hence finally $x \in (1, \infty)$.

Example 6

Find the domain of the function $f(x) = e^{x^2-1}$.

Solution
The exponential function e^x is defined for all the real numbers. Hence the domain of f depends only on the exponent. The exponent is a polynomial, so its domain is also the set of real numbers. It means that $D(f) = R$.

Example 7
Find the domain of the function $f(x) = 1/(2^x - 8)$.

Solution
The expression 2^x is defined for all the real numbers. Because of that the only thing that matters is the fact that it is forbidden to divide by 0, and thus $2^x \neq 8$. As $8 = 2^3$, the inequality takes the form $2^x \neq 2^3$, and finally $x \neq 3$. Therefore $D(f) = R \setminus \{3\}$.

EXERCISES

1. Simplify the expressions:
   a) $(1/4)^x \cdot 2^x/(1/8)^x$;
   b) $9^x \cdot 8^x/6^{2x}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) $(7)^{5x}\cdot(49)^{3x+2}/(7)^{5x+4}$;
d) $(2)^{2x}\cdot(4)^{4x+2}/(8)^{x-8}$;
e) $(1/4)^{x}\cdot(1/8)^{3x+2}/(1/16)^{x-8}$.

2. Solve the equations and inequalities:
   a) $2^x > 8$;
   b) $(1/2)^{3x+1} = (1/4)^{5x-10}$;
   c) $(1/9)^{2x+5} < (1/27)^{3x-2}$;
   d) $4^{2x-5} = 8^{3x-1}$;
   e) $2\cdot 3^{2x} < 3\cdot 4^{x}$.
3. Find the domains of the functions:
   a) $f(x) = 1 / 2^{2x}$;
   b) $f(x) = \sqrt{7^{2x} - 49}$;
   c) $f(x) = 1/\sqrt{27 - 3^{2x}}$;
   d) $f(x) = \dfrac{1}{e^{x^2-3x+2} - 1}$;
   e) $f(x) = (e^{2x} + 1) / (e^{2x} - 1)$.

SOLUTIONS

1. Sample answers:
   a) 16^x; b) 2^x; c) 7^{6x};
   d) 2^{7x+28}; e) $(1/2)^{7x+38}$.
2. a) $x \in (3, \infty)$; b) $x = 3$; c) $x \in (-\infty, 16/5)$;
   d) $x = -7/5$; e) $x \in (-\infty, 1/2)$.
3. a) $D(f) = R$; b) $D(f) = [1, \infty)$; c) $D(f) = (-\infty, 3/2)$;
   d) $D(f) = R \setminus \{1, 2\}$; e) $D(f) = R \setminus \{0\}$.

6.4. LOGARITHMIC FUNCTION

THEORY IN A NUTSHELL

Logarithmic function is the function $f(x) = \log_a x$ (we read it: base a logarithm of x), where a is parameter. We assume that $a \in (0, 1) \cup (1, \infty)$. It is the **inverse function**[17] of the exponential function: $y = \log_a x$ if and only if $x = a^y$, for example $\log_2 8 = 3$, because $2^3 = 8$.

The logarithmic function is defined only for $x > 0$. Its behavior depends on the value of a. If $a \in (0, 1)$, the function is decreasing: for x approaching 0, its value increases very rapidly to ∞, and for x increasing to ∞, it decreases to $-\infty$. When $a > 0$, the function is increasing: for x approaching 0 its value decreases

[17] If you do not remember what that means, recall the definition of inverse relations from Section 3.1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

rapidly to $-\infty$, and for x increasing to ∞, it increases to ∞. Regardless of the value of a, the logarithmic function has always one zero (i.e., intersection with the axis $0x$), at $x = 1$: $f(1) = \log_a 1 = 0$. Sample behavior of the logarithmic function for various values of a has been illustrated in Fig. 6.8.

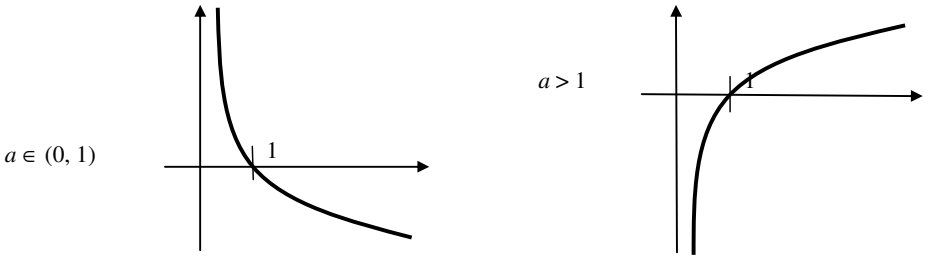


Fig. 6.8. The logarithmic function

Basic properties of logarithms:

1. Operations on logarithms having the same base:
   $\log_a(xy) = \log_a x + \log_a y$, $\log_a(x/y) = \log_a x - \log_a y$.
2. Logarithm of a power: $\log_a(x^y) = y \log_a x$.
3. Change-of-base formula: $(\log_a x)/(\log_a y) = \log_y x$.
4. Composition of logarithmic and exponential functions: $\log_a(a^x) = a^{\log_a x} = x$.

To solve a **logarithmic equation**, you must transform both sides to the form of logarithms having the same base, and then compare the arguments. In the case of a **logarithmic inequality** we proceed similarly, bearing in mind the fact that it is necessary to reverse the inequality, when $a < 1$ (note that the function decreases in this case).

From the point of view of applications, in the remainder of the book we will be particularly interested in the so-called **natural logarithm**, denoted by $\ln(x)$, being logarithm to the base e.

Note: the notation $\ln^2 x$ is equivalent to the notation $(\ln x)^2$. The notation $\ln x^2$, in turn, is equivalent to the notation $\ln(x^2)$ (note the order of operations).

In order to implement the logarithms in **WolframAlpha®**, you can use the notation "log_a(x)" for general logarithm and "ln(x)" for the natural one. For example, if you want to solve the equation from Example 5 below, you should type:

"solve ln(x-1)+ln(x-2)=ln(6)".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

If you want to find the domain of the function $f(x) = 1 / \ln x$, you can type:

"domain f(x)=1/ln(x)".

EXAMPLES

Example 1
Simplify the expression: $\ln x - \ln (x^2 + 2)$.

Solution
We deal with the logarithms to the same base (e). Because of that we use property 1: $\ln x - \ln (x^2 + 2) = \ln [x / (x^2 + 2)]$.

Example 2
Move the exponent outside the logarithm: $\log_2 x^5$.

Solution
We use the second property to obtain: $\log_2 x^5 = 5 \cdot \log_2 x$.

Example 3
Transform the expression in such a way that there will be only natural logarithms in the final form: $\log_{(2x-3)}(x^2 - 2x)$.

Solution
It is necessary to use the change-of-base formula (property 3):

$\log_{(2x-3)}(x^2 - 2x) = \ln (x^2 - 2x) / \ln (2x - 3)$.

Example 4
Simplify the expression: $\ln (e^{2x})$.

Solution
As we deal with the composition of the logarithmic and exponential functions, we use property 4: $\ln (e^{2x}) = 2x$.

Example 5
Solve the equation: $\ln (x - 1) + \ln (x - 2) = \ln 6$.

Solution
As the argument of a logarithmic function must be positive, we must make the following assumptions: $x - 1 > 0$ and $x - 2 > 0$. It follows that $x > 2$. Now we

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

move to the main part of the problem. We begin with transforming the left side. By using property 1 we obtain:

$\ln [(x - 1)(x - 2)] = \ln 6,$
$\ln (x^2 - 3x + 2) = \ln 6.$

As on both sides there are logarithms having the same base, we can compare the arguments:

$x^2 - 3x + 2 = 6.$

It leads us to the solution: $x = -1 \vee x = 4$. As the condition $x > 2$ must be satisfied, the final answer is $x = 4$.

Example 6
Solve the inequality: $\ln (2x^2 - x) \geq 0$.

Solution

Just like before, we have to make some assumptions: $2x^2 - x > 0$, hence $x \in (-\infty, 0) \cup (1/2, \infty)$.

Using property 4, we obtain: $0 = \ln e^0 = \ln 1$. Thus the inequality may be written in the form:

$\ln (2x^2 - x) \geq \ln 1.$

As both logarithms have the same base (e), we can compare the arguments. We do not reverse the inequality, because $e > 1$ (recall that $e = 2.71828\ldots$). Thus we obtain:

$2x^2 - x \geq 1,$

i.e. $x \in (-\infty, -1/2] \cup [1, \infty)$. Taking into account the assumption, we finally obtain $x \in (-\infty, -1/2] \cup [1, \infty)$.

Example 7
Find the domain of the function $f(x) = \ln(x^2 - 2x)$.

Solution
Logarithmic function is defined only for positive arguments. For that reason the condition $x^2 - 2x > 0$ must be satisfied. Hence $D(f) = (-\infty, 0) \cup (2, \infty)$.

Example 8
Find the domain of the function $f(x) = \ln(\ln(3x - 2x^2))$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

This time we deal with a composition of two logarithmic functions. For that reason two conditions must be satisfied: $3x - 2x^2 > 0$ and $\ln(3x - 2x^2) > 0$. The solution of the first one is (0, 3/2), and the solution of the second one is (1/2, 1). The domain is the intersection of the above sets, hence $D(f) = (1/2, 1)$.

EXERCISES

1. Simplify the expressions:
   a) $\ln(2x) - \ln(x + 2)$;
   b) $\ln x + \ln(x^2 - 7)$;
   c) $\ln(x + 2) + \ln(x^2 + 2) - \ln(x^3 + 2)$;
   d) $\ln e^{2x + 5}$;
   e) $\ln e^{x - 17}$;
   f) $e^{\ln(7x + 17)}$.
2. Move the exponent outside the logarithm:
   a) $\ln x^7$;
   b) $\ln(x + 2)^{17}$.
3. Move the exponent inside the logarithm:
   a) $12\ln(x - 3)$;
   b) $5\ln(2x + 5)$.
4. Express using the natural logarithms:
   a) $\log_2(x - 7)$;
   b) $\log_{(2x - 3)}(3x + 11)$;
   c) $\log_{(x + 2)}(x^2 - 7)$.
5. Solve the equations and inequalities:
   a) $\ln(x^2) = 3$;
   b) $\ln(x^2 + 1) + \ln(x + 2) = \ln 6$;
   c) $\ln(x^2 - 1) > 0$;
   d) $2\ln x < 1$;
   e) $\ln x + \ln(2x) - \ln(3x) = 1$.
6. Find the domains of the functions:
   a) $\ln(\ln(\ln x))$;
   b) $\ln[(x - 1)^2]$;
   c) $2\ln(x - 1)$;
   d) $\ln(x - 1) / \ln(x - 2)$;
   e) $\sqrt{1 - \ln^2 x}$.

SOLUTIONS

1. a) $\ln[2x / (x + 2)]$; b) $\ln[x(x^2 - 7)]$;
   c) $\ln[(x + 2)(x^2 + 2) / (x^3 + 2)]$; d) $2x + 5$;
   e) $x - 17$; f) $7x + 17$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. a) $7\ln x$; b) $17\ln(x+2)$.

3. a) $\ln(x-3)^{12}$; b) $\ln(2x+5)^{5}$.

4. a) $\ln(x-7)/\ln 2$; b) $\ln(3x+11)/\ln(2x-3)$;
   c) $\ln(x^2-7)/\ln(x+2)$.

5. a) $x = -e^{3/2} \vee x = e^{3/2}$; b) $x = 1$;
   c) $x \in (-\infty, -\sqrt{2}) \cup (\sqrt{2}, \infty)$; d) $x \in (-\infty, e^{1/2})$;
   e) $x = 3/2e$.

6. a) $D(f) = (e, \infty)$; b) $D(f) = R \setminus \{1\}$;
   c) $D(f) = (1, \infty)$; d) $D(f) = (2, 3) \cup (3, \infty)$;
   e) $D(f) = [1/e, e]$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 7
SEQUENCES AND SERIES

7.1. SEQUENCES AND THEIR LIMITS

THEORY IN A NUTSHELL

Recall that sequence is a function whose domain is a subset of the set of natural numbers[18]. We usually denote a sequence by (a_n), and its n-th element by a_n. The sequences may be finite or infinite. Example of a finite sequence is (0, 1, 3, 7, 15). Example of an infinite sequence is the sequence described with the formula $a_n = 2^n - 1,\ n \in N$, i.e. the sequence consisting of the elements: $a_0 = 2^0 - 1 = 0,\ a_1 = 2^1 - 1 = 1,\ a_2 = 2^2 - 1 = 3,\ a_3 = 2^3 - 1 = 7, \ldots$ (it is not the same sequence as the firs one, in this case we can list the elements indefinitely).
From our point of view an important concept is the **limit of a sequence**, which is denoted by: $\lim_{n\to\infty} a_n$, which we read as „the limit of a_n as n approaches infinity".
We can also write it shorter as: $\lim a_n$ (we read it „the limit of a_n"). It is the value, to which the elements of a sequence approach as n increases indefinitely. It may happen that the values are increasing or decreasing over in unlimited way and then we say that the limit is respectively ∞ or $-\infty$. It may also happen that the limit does not exist.
You will not need the formal definition of limit, but it may be useful to know its graphical interpretation, illustrated in Fig. 7.1 (note that the plot is a set of points and not a continuous line, as is the case of most of the functions).

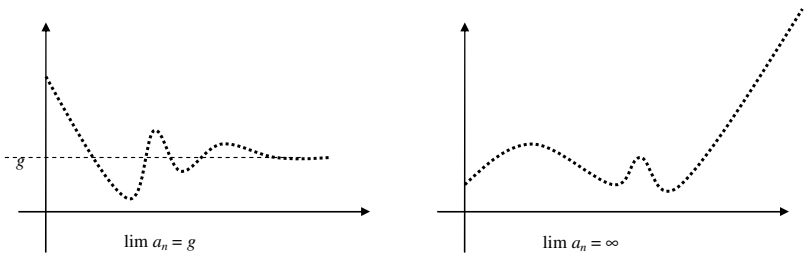


Fig. 7.1. Limit of a sequence

For the calculation of limits of sequences you will need few facts about them.

[18] Section 3.2, Definition 3.4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

1. Limit of a power function: $\lim_{n\to\infty} n^a = \infty$, when $a > 0$, $\lim_{n\to\infty} n^a = 0$, when $a < 0$ (for example $\lim_{n\to\infty}(1/n) = 0$).
2. Limit of an exponential function: $\lim_{n\to\infty} a^n = \infty$, when $a > 1$, $\lim_{n\to\infty} a^n = 0$, when $a \in (0, 1)$.
3. Number e: $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$. Note: instead of n you can use in this formula any expression that approaches either ∞ or –∞.
4. Operations on limits:

$$\lim(a_n + b_n) = \lim a_n + \lim b_n, \quad \lim(a_n - b_n) = \lim a_n - \lim b_n,$$

$$\lim(a_n b_n) = (\lim a_n)(\lim b_n), \quad \lim(a_n / b_n) = (\lim a_n)/(\lim b_n),$$

$$\lim(c + a_n) = c + \lim a_n, \quad \lim(c \cdot a_n) = c \lim a_n.$$

A very important thing is the ability to perform operations on expressions including infinity and zero. The basic rules can be written in the following simplified way ("0^+" means an expression tending to 0 from above, "0^-" from below; a is a constant different from 0):

$[a \cdot \infty] = \infty$, $[\infty + \infty] = \infty$, $[\infty \cdot \infty] = \infty$, $[1/0^+] = \infty$, $[1/0^-] = -\infty$, $[1/\infty] = 0$, $[\infty^a] = \infty$, gdy $a > 0$, $[\infty^a] = 0$, gdy $a < 0$, $[\infty^\infty] = \infty$.

Sometimes, when calculating the value of a limit you will come across the so-called **indeterminate form**. There are seven expressions of this type:

[0/0], [∞/∞], [0 · ∞], [∞ – ∞], [∞^0], [1^∞] and [0^0].

If there is an indeterminate form, it can not be predicted unambiguously what is the value of the limit. For that reason you must convert the expression whose limit is calculated to a more convenient form. Note that indeterminate forms are written in square brackets. Use this notation also in the case of all the indirect, auxiliary calculations, which need not be formally correct, but will help you to find the final value of the limit, for example [∞·∞], [∞^2] or [$1/0^+$].
Sometimes the transformations are not enough – then the following theorem may be helpful.

Theorem 7.1. The Squeeze Theorem
If $\lim_{n\to\infty} b_n = g$, $\lim_{n\to\infty} c_n = g$ and $\exists_M \forall_{n>M}\ b_n \le a_n \le c_n$, then $\lim_{n\to\infty} a_n = g$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

This theorem allows to find the limit of a sequence (a_n) even if its formula is quite complicated, but we can find two sequences of simpler structure, satisfying two conditions: first, from a certain point (not necessarily for all elements), the elements of one of them are not less a second not greater than the elements of (a_n); secondly, the two sequences must have the same limit. Graphical interpretation of the theorem is presented in Fig. 7.2.

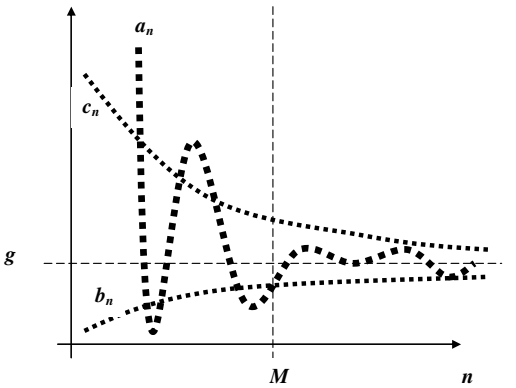


Fig. 7.2. The Squeeze Theorem

If you want to find the limit of sequence (a_n) using **WolframAlpha®**, you should use the command "limit". For example, if you want to calculate the limit of the sequence from Example 1 below, you can type:

"limit 7n^3 - 200 n^2 - 510 as n-> infinity".

It is also possible to use abbreviations of the words "limit" and "infinity":

"lim 7n^3 - 200 n^2 - 510 as n-> inf".

EXAMPLES

Example 1
Find the limit of the sequence:

$a_n = 7n^3 - 200n^2 - 510$.

Solution
In case of the polynomials we begin with factoring out the highest power of n:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim a_n = \lim_{n\to\infty}\left(7n^3 - 200n^2 - 510\right) = \lim_{n\to\infty} n^3\left(7 - 200/n - 510/n^3\right) = \ldots$$

From the first property of limits it follows that $200/n$, as well as $510/n^3$, tend to 0. Thus we can proceed:

$$\ldots = \lim_{n\to\infty} n^3\left(7 - 0 - 0\right) = \left[7 \cdot \infty^3\right] = \infty .$$

The last equality follows from the above properties of operations on the expressions with infinity. Note the square brackets, used together with the expression "$7 \cdot \infty^3$", which is not correct from the formal point of view, but allows us to save the intermediate calculations.

Example 2
Find the limit of the sequence:

$$a_n = \frac{n^3 + 2n^2 + 5}{3n^3 - 4n} .$$

Solution
In case of the rational functions, similiarly as in case of the polynomials we begin with factoring out the highest power of n:

$$\lim a_n = \lim_{n\to\infty} \frac{n^3 + 2n^2 + 5}{3n^3 - 4n} = \lim_{n\to\infty} \frac{n^3\left(1 + 2/n + 5/n^3\right)}{n^3\left(3 - 4/n^2\right)} = \ldots$$

Similarly as above we use the fact that some summands approach 0. After substituting 0 to respective places and simplifying the fraction we obtain:

$$\ldots = \lim_{n\to\infty} \frac{n^3\left(1 + 0 + 0\right)}{n^3\left(3 - 0\right)} = 1/3 .$$

Example 3
Find the limit of the sequence:

$$a_n = \frac{\sqrt{n^3 + 2n^2 + 3}}{3n + 2} .$$

Solution
We act like before – we factor out the highest powers in numerator and denominator:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim a_n = \lim_{n\to\infty} \frac{\sqrt{n^3\left(1+2/n+3/n^3\right)}}{n\left(3+2/n\right)} = \lim_{n\to\infty} \frac{\sqrt{n^3\left(1+0+0\right)}}{n\left(3+0\right)} = \lim_{n\to\infty} \frac{\sqrt{n^3}}{3n} = \lim_{n\to\infty} \frac{n^{3/2}}{3n} =$$

$$= \lim_{n\to\infty} \frac{n^{1/2}}{3} = \infty .$$

The last equality follows from the fact that n^a tends to ∞ when $a > 0$.

Example 4
Find the limit of the sequence:

$$a_n = \frac{5\cdot 2^n + 4\cdot 3^n}{5\cdot 3^{n+1} - 7\cdot 2^n} .$$

Solution
Also in this case we factor out the highest power in numerator and denominator. This time we deal with the powers having the same exponent, so we choose the highest base. We also use the fact that $3^{n+1} = 3^n \cdot 3^1 = 3 \cdot 3^n$.

$$\lim a_n = \lim_{n\to\infty} \frac{3^n(5\cdot(2/3)^n + 4)}{3^n(5\cdot 3 - 7\cdot(2/3)^n)} = \ldots$$

As a^n tends to 0 when $a < 1$, we can skip the expressions of the form $(2/3)^n$. By simplifying the fraction we obtain:

$$\ldots = \lim_{n\to\infty} \frac{3^n(0+4)}{3^n(5\cdot 3 - 0)} = 4/15 .$$

Example 5
Find the limit of the sequence:

$$a_n = \sqrt[n]{3\cdot 5^n + 5\cdot 7^n - 100\cdot 2^n} .$$

Solution
Also in this case we factor out the highest power.

$$\lim a_n = \lim_{n\to\infty} \sqrt[n]{7^n(3\cdot(5/7)^n + 5 - 100\cdot(2/7)^n)} =$$

$$= \lim_{n\to\infty} \sqrt[n]{7^n(0+5-0)} = \lim_{n\to\infty} \sqrt[n]{7^n\cdot 5} = \ldots$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The n-th root can be written as a power with the exponent $1/n$. By performing standard operations on powers[19] and bearing in mind that $1/n$ tends to 0, we obtain finally:

$$... = \lim_{n\to\infty}\left(7^n \cdot 5\right)^{1/n} = \lim_{n\to\infty}\left[\left(7^n\right)^{1/n} \cdot 5^{1/n}\right] = \lim_{n\to\infty}\left(7^{n\cdot(1/n)} \cdot 5^{1/n}\right) = 7^1 \cdot 5^0 = 7 \cdot 1 = 7\,.$$

Example 6
Find the limit of the sequence:

$$a_n = \frac{\sqrt{2n^2+2n+5} - \sqrt{n^2+3}}{5}\,.$$

Solution
Again, we factor out the highest powers:

$$\lim a_n = \lim_{n\to\infty}\frac{\sqrt{n^2(2+2/n+5/n^2)} - \sqrt{n^2(1+3/n^2)}}{5} = \lim_{n\to\infty}\frac{\sqrt{2n^2} - \sqrt{n^2}}{5} = ...$$

By writing the root in the form of a power and using standard operations we obtain:

$$... = \lim_{n\to\infty}\frac{n(\sqrt{2}-1)}{5} = \infty\,.$$

Example 7
Find the limit of the sequence:

$$a_n = \frac{\sqrt{n^2+2n+5} - \sqrt{n^2+3}}{5}\,.$$

Solution
Proceeding exactly as in the previous example, you will not get a solution:

$$\lim a_n = \lim_{n\to\infty}\frac{\sqrt{n^2+2n+5} - \sqrt{n^2+3}}{5} = ... = \lim_{n\to\infty}\frac{n(1-1)}{5} = [\infty \cdot 0]\,.$$

[19] If you do not remember those operations, go back to Section 6.3.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that as a result of similar actions, this time we obtained one of the seven indeterminate forms. So we have to use a different method to determine the limit. This method consists in multiplying and dividing the expression with roots by its conjugate, i.e. the expression, in which we replace the sign "–" between both parts with the sign "+":

$$\lim a_n = \lim_{n\to\infty}\left(\frac{\sqrt{n^2+2n+5}-\sqrt{n^2+3}}{5}\cdot\frac{\sqrt{n^2+2n+5}+\sqrt{n^2+3}}{\sqrt{n^2+2n+5}+\sqrt{n^2+3}}\right)=\ldots$$

Why do we multiply and divide by the same – it does not change the value of the whole expression? And that is correct. The value of the expression cannot be changed, but the expression in a new form can be easily transformed. In the numerator we will ue one of the short multiplication formulae, and then we will factor out the highest power as in the previous examples:

$$\ldots=\lim_{n\to\infty}\left(\frac{(n^2+2n+5)-(n^2+3)}{5(\sqrt{n^2(1+2/n+5/n^2)}+\sqrt{n^2(1+3/n^2)})}\right)=\lim_{n\to\infty}\left(\frac{2n+2}{5(\sqrt{n^2}+\sqrt{n^2})}\right)=$$

$$=\lim_{n\to\infty}\left(\frac{n(2+2/n)}{5\cdot 2n}\right)=2/10=1/5.$$

Note what is the main difference between the last two examples. In the first one, the coefficients by the highest powers of n (i.e., n^2) are different (2 and 1), and therefore it was possible to find the limit by using the standard factoring out. In the second case these coefficients are equal, and therefore it was necessary to multiply and divide by the conjugate of the numerator.

Example 8
Find the limit of the sequence:

$$a_n=\left(\frac{3n+2}{4n+7}\right)^{2n}.$$

Solution
Also in this case we factor out the highest power of n (i.e., n) (independently in the numerator and denominator), and then raise the result to a power:

$$\lim a_n=\lim_{n\to\infty}\left(\frac{n(3+2/n)}{n(4+7/n)}\right)^{2n}=\lim_{n\to\infty}(3/4)^{2n}=0\,.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 9
Find the limit of the sequence:

$$a_n = \left(\frac{3n+2}{3n+7}\right)^{2n}.$$

Solution
The method from the last example will not be useful this time:

$$\lim a_n = \lim_{n\to\infty}\left(\frac{n(3+2/n)}{n(3+7/n)}\right)^{2n} = \lim_{n\to\infty}(3/3)^{2n} = [1^{\infty}].$$

As you can see, again, the result is one of the seven indeterminate forms, so you must use a different method. We start with almost the same, with the difference that we exclude the highest power of n together with the coefficient, and then separately raise numerator and denominator to the respective power:

$$\lim a_n = \lim_{n\to\infty}\frac{\left[3n\left(1+\frac{2}{3n}\right)\right]^{2n}}{\left[3n\left(1+\frac{7}{3n}\right)\right]^{2n}} = \lim_{n\to\infty}\frac{\left(1+\frac{2}{3n}\right)^{2n}}{\left(1+\frac{7}{3n}\right)^{2n}} = \ldots$$

This is very similar to the third basic type of limits. But the number in the exponent should be the reciprocal of the expression added to 1 in the brackets. For example the expression in the numerator is $2/(3n)$, hence the exponent should be $3n/2$. Thus it is necessary to write the current exponent $(2n)$ in such a way that the factor $3n/2$ is a part of it. It is necessary to multiply the last expression by $4/3$: $2n = 3n/2 \cdot 4/3$. Where this number comes from? We just divide the current exponent $(2n)$ by the one that we want to obtain $(3n/2)$: $2n/(3n/2) = 4/3$. Similarly there is $7/(3n)$ in the denominator, so we want to factor out $3n/7$, and thus we write $2n$ as $3n/7 \cdot 14/3$. Using the standard properties of powers we obtain:

$$\ldots = \lim_{n\to\infty}\frac{\left(1+\frac{2}{3n}\right)^{3n/2\cdot 4/3}}{\left(1+\frac{7}{3n}\right)^{3n/7\cdot 14/3}} = \lim_{n\to\infty}\frac{\left[\left(1+\frac{2}{3n}\right)^{3n/2}\right]^{4/3}}{\left[\left(1+\frac{7}{3n}\right)^{3n/7}\right]^{14/3}} = \ldots$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

According to the third property of limits, both expressions in the square brackets tend to e. Using the formula for dividing the powers of the same base we finally obtain:

$$\ldots = \frac{e^{4/3}}{e^{14/3}} = e^{-10/3}.$$

Note what is the main difference between the last two examples. In the first one, the coefficients by the highest powers of n (i.e., n) are different (3 in the numerator and 4 in the denominator), and therefore it was possible to find the limit by using the standard factoring out. In the second case these coefficients are equal, and therefore it was necessary to use the formula for e.

Example 10
Find the limit of the sequence, using the Squeeze Theorem:

$$a_n = \frac{2n + 5 \cdot (-1)^n}{5n + 3 \cdot (-1)^n}.$$

Solution
The element that makes the standard method useless is $(-1)^n$ – this expression does not have a limit, taking alternately the values –1 and 1. For that reason it is necessary to find its lower and upper bounds. The best one are –1 and 1: $(-1)^n \le 1$ and $(-1)^n \ge -1$ for each value of n. In order to bound the whole term of the sequence we must remember that in order to increase the value of a positive valued fraction, we increase its numerator and decrease denominator. In order to decrease the value of the fraction, we decrease the numerator and increase the denominator. This way we obtain the lower (b_n) and upper (c_n) bound of a_n:

$$a_n = \frac{2n + 5 \cdot (-1)^n}{5n + 3 \cdot (-1)^n} \ge \frac{2n + 5 \cdot (-1)}{5n + 3 \cdot 1} = \frac{2n - 5}{5n + 3} = b_n,$$

$$a_n = \frac{2n + 5 \cdot (-1)^n}{5n + 3 \cdot (-1)^n} \le \frac{2n + 5 \cdot 1}{5n + 3 \cdot (-1)} = \frac{2n + 5}{5n - 3} = c_n.$$

We find the limits of both bounds:

$$\lim b_n = \lim_{n \to \infty} \frac{2n - 5}{5n + 3} = \lim_{n \to \infty} \frac{n(2 - 5/n)}{n(5 + 3/n)} = \frac{2}{5},$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim c_n = \lim_{n\to\infty} \frac{2n+5}{5n-3} = \lim_{n\to\infty} \frac{n(2+5/n)}{n(5-3/n)} = \frac{2}{5}.$$

Since $b_n \le a_n \le c_n$, and the limits of both bounds are equal: $\lim b_n = \lim c_n = 2/5$, also the limit of a_n has the same value: $\lim a_n = 2/5$.

Example 11
Find the limit of the sequence, using the Squeeze Theorem:

$$a_n = \sqrt[n]{5\cdot 7^n + 3\cdot 3^n + 4\cdot(-5)^n}.$$

Solution
This time the component $(-5)^n$ makes trouble: every second term is equal to 5^n (and tends to ∞), and the other ones are equal to -5^n (and approach $-\infty$). For that reason we will use the inequalities: $(-5)^n \ge -5^n$ and $(-5)^n \le 5^n$, and the bounds will be:

$$a_n = \sqrt[n]{5\cdot 7^n + 3\cdot 3^n + 4\cdot(-5)^n} \ge \sqrt[n]{5\cdot 7^n + 3\cdot 3^n - 4\cdot 5^n} = b_n,$$

$$a_n = \sqrt[n]{5\cdot 7^n + 3\cdot 3^n + 4\cdot(-5)^n} \le \sqrt[n]{5\cdot 7^n + 3\cdot 3^n + 4\cdot 5^n} = c_n.$$

We find the limits of both bounds:

$$\lim b_n = \lim_{n\to\infty} \sqrt[n]{5\cdot 7^n + 3\cdot 3^n - 4\cdot 5^n} = \lim_{n\to\infty} \sqrt[n]{7^n(5 + 3\cdot(3/7)^n - 4\cdot(5/7)^n)} = 7,$$

$$\lim c_n = \lim_{n\to\infty} \sqrt[n]{5\cdot 7^n + 3\cdot 3^n + 4\cdot 5^n} = \lim_{n\to\infty} \sqrt[n]{7^n(5 + 3\cdot(3/7)^n + 4\cdot(5/7)^n)} = 7.$$

Since $b_n \le a_n \le c_n$, and the limits of both bounds are equal: $\lim b_n = \lim c_n = 7$, also the limit of a_n has the same value: $\lim a_n = 7$.

Example 12
Find the limit of the sequence, using the Squeeze Theorem:

$$a_n = \frac{n^2+1}{n^3+1} + \frac{n^2+1}{n^3+2} + \ldots + \frac{n^2+1}{n^3+n}.$$

Solution
Each term of the sequence is a sum of n fractions having different denominators. And this is a problem – in order to sum up the fractions we should find

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

a common denominator, for example by multiplying the denominators of all the fractions. However, it is impossible in this case – for each term of the sequence the number of fractions is different. In order to find proper bounds we will change slightly the denominators of the fractions. In order to decrease the value of a term we will equate the denominators with the biggest one (i.e., $n^2 + n$). In order to increase the value of a term we will equate the denominators with the smallest one (i.e., $n^2 + 1$).

$$a_n = \frac{n^2+1}{n^3+1} + \frac{n^2+1}{n^3+2} + \ldots + \frac{n^2+1}{n^3+n} \geq \frac{n^2+1}{n^3+n} + \frac{n^2+1}{n^3+n} + \ldots + \frac{n^2+1}{n^3+n} = n \cdot \frac{n^2+1}{n^3+n} = b_n,$$

$$a_n = \frac{n^2+1}{n^3+1} + \frac{n^2+1}{n^3+2} + \ldots + \frac{n^2+1}{n^3+n} \leq \frac{n^2+1}{n^3+1} + \frac{n^2+1}{n^3+1} + \ldots + \frac{n^2+1}{n^3+1} = n \cdot \frac{n^2+1}{n^3+1} = c_n.$$

We found the limits of both bounds:

$$\lim b_n = \lim_{n\to\infty} n \cdot \frac{n^2+1}{n^3+n} = \lim_{n\to\infty} n \cdot \frac{n^2(1+1/n^2)}{n^3(1+1/n^2)} = 1,$$

$$\lim b_n = \lim_{n\to\infty} n \cdot \frac{n^2+1}{n^3+1} = \lim_{n\to\infty} n \cdot \frac{n^2(1+1/n^2)}{n^3(1+1/n^3)} = 1.$$

Since $b_n \leq a_n \leq c_n$, and the limits of both bounds are equal: $\lim b_n = \lim c_n = 1$, also the limit of a_n has the same value: $\lim a_n = 1$.

EXERCISES

1. Find the limits of the sequences, using standard definitions and transformations:
   a) $a_n = 17n^4 + 20n - 666$;
   b) $a_n = \frac{2n^3 - 7n^2 + 5}{n^3 + 4n}$;
   c) $a_n = \frac{n^2 + 2n - 1}{7n^4 + 4n}$;
   d) $a_n = \frac{\sqrt{n^4 - 2n + 7}}{11n^2 - 5}$;
   e) $a_n = \frac{5 \cdot 3^n + 11 \cdot 7^n}{3 \cdot 7^{n+1} - 17 \cdot 2^n}$;
   f) $a_n = \frac{2 \cdot 6^n + 6 \cdot 5^n}{7 \cdot 3^{n+1} + 11 \cdot 2^n}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

g) $a_n = \sqrt[n]{2 \cdot 6^n + 5 \cdot 5^n - 10 \cdot 4^n}$;

h) $a_n = \sqrt[n]{7 \cdot 5^n + 5 \cdot 7^n - 6 \cdot 6^n}$;

i) $a_n = \dfrac{\sqrt{n^2 + 2n + 5} - \sqrt{2n^2 + 7}}{8}$;

j) $a_n = \dfrac{\sqrt{n^2 + n + 3} - \sqrt{n^2 + 2}}{6}$;

k) $a_n = \dfrac{\sqrt{n^3 + 7n + 5} - \sqrt{n^3 + 3}}{17}$;

l) $a_n = \left(\dfrac{8n - 2}{5n - 7} \right)^{3n}$;

m) $a_n = \left(\dfrac{11n + 9}{11n + 7} \right)^{4n}$;

n) $a_n = \left(\dfrac{13n - 3}{13n - 5} \right)^{7n}$.

2. Find the limits of the sequences, using the Squeeze Theorem:

a) $a_n = \dfrac{5n + 4 \cdot (-1)^n}{7n + 2 \cdot (-1)^n}$;

b) $a_n = \dfrac{11n + 3 \cdot (-1)^n}{2n + 7 \cdot (-1)^n}$;

c) $a_n = \sqrt[n]{6 \cdot 8^n + 4 \cdot 4^n + 5 \cdot (-6)^n}$;

d) $a_n = \sqrt[n]{7 \cdot 9^n + 5 \cdot 5^n + 6 \cdot (-7)^n}$;

e) $a_n = \dfrac{n}{n^2 + 1} + \dfrac{n}{n^2 + 2} + \ldots + \dfrac{n}{n^2 + n}$;

f) $a_n = \dfrac{n^3}{n^4 + 1} + \dfrac{n^3}{n^4 + 2} + \ldots + \dfrac{n^3}{n^4 + n}$.

SOLUTIONS

1. a) ∞; b) 2; c) 0; d) 1/11;
   e) 11/21; f) ∞; g) 6; h) 7;
   i) –∞; j) 1/12; k) 0; l): ∞;
   m) $e^{8/11}$; n) $e^{14/13}$.
2. a) 5/7; b) 11/2; c) 8; d) 9;
   e) 1; f) 1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

7.2. ARITHMETIC AND GEOMETRIC PROGRESSIONS

THEORY IN A NUTSHELL

Arithmetic progression (or **arithmetic sequence**) is a sequence in which the difference between any two consecutive terms is constant, i.e. $a_{n+1} - a_n = r$. The number r is called **the difference of the progression**. The formula for the n-th term of arithmetic progression is:

$a_n = a_1 + (n-1)r.$

The sum of n first terms of arithmetic progression is equal to:

$$S_n = \frac{a_1 + a_n}{2} n .$$

Geometric progression (geometric sequence) is a sequence in which the quotient of two consecutive terms is constant, i.e. $a_{n+1} / a_n = q$. The number q is called the **common ratio of the progression**. The formula for the n-th term of geometric progression is:

$a_n = a_1 q^{n-1}.$

The sum of n first terms of geometric progression is equal to:

$$S_n = a_1 \frac{1-q^n}{1-q} .$$

In the case of a geometric sequence it is also possible to find the sum even if that sequence is infinite, provided that one additional condition is satisfied:

$$S_\infty = \frac{a_1}{1-q}, \quad |q| < 1 .$$

In order to calculate the sum of an arithmetic or geometric progression with **WolframAlpha®**, you can just type few first terms of the progression (at least three) together with the "+" symbols and the last term after dots, if it is a finite sequence. For example, if you want to solve Example 1 below, you should type:

"1+4+7+...+100".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

If you want to find the sum of the infinite sequence from Example 3 below, you can type:

"4+8/3+16/9+...".

EXAMPLES

Example 1
Find the sum: 1 + 4 + 7 + … + 100.

Solution
We deal with an arithmetic progression whose difference equals to 4 – 1 = 7 – 4 = 3, the first term is 1, and the last term is 100. By transforming the formula for the n-th term $a_n = a_1 + (n-1)r$ to the form $n = (a_n - a_1) / r + 1$, we obtain that the number of terms is equal to $n = (100 - 1)/3 + 1 = 34$. Hence the sum of these 34 terms equals to:

$S_{34} = (1 + 100)/2 \cdot 34 = 1717$.

Example 2
Find the sum:

$$4 + \frac{8}{3} + \frac{16}{9} + \frac{32}{27} + \frac{64}{81}.$$

Solution
This time we are dealing with a geometric progression with common ratio equal to 8/3 / 4 = (16/9) / (8/3) = 2/3, and the firs term 4. Since the sequence consists of five terms, the sum is equal to:

$$S_5 = 4 \cdot \frac{1-(2/3)^n}{1-2/3} = 844/81.$$

Example 3

Find the sum: $4 + \frac{8}{3} + \frac{16}{9} + \frac{32}{27} + \frac{64}{81} + ...$

Solution
Just as in the previous example we are dealing with a geometric progression (the common ratio and the first term are the same as before). This time it is an infinite sequence. The common ratio satisfies the condition $|2/3| < 1$, so we can find the sum:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$S_\infty = \frac{4}{1-2/3} = 12 .$$

EXERCISES

1. Find the sums of arithmetic progressions:
   a) $2+4+6+...+50$;
   b) $1+3+5+...+103$;
   c) $5+11+17+...+101$.
2. Find the sums of geometric progressions:
   a) $2+\frac{2}{5}+\frac{2}{25}+\frac{2}{125}+\frac{2}{625}$;
   b) $\frac{100}{1{,}1}+\frac{100}{1{,}1^2}+...+\frac{100}{1{,}1^{10}}$;
   c) $2+\frac{2}{5}+\frac{2}{25}+\frac{2}{125}+\frac{2}{625}+...$;
   d) $\frac{100}{1{,}1}+\frac{100}{1{,}1^2}+\frac{100}{1{,}1^3}+\frac{100}{1{,}1^4}+...$;
   e) $\frac{100}{1{,}1}+\frac{100\cdot 1{,}05}{1{,}1^2}+\frac{100\cdot 1{,}05^2}{1{,}1^3}+\frac{100\cdot 1{,}05^3}{1{,}1^4}+...$

SOLUTIONS

1. a) 650; b) 2704; c) 901.
2. a) 1562/625; b) $1000(1 - 1/1{,}1^{10}) \approx 614{,}46$;
   c) 5/2; d) 1000; e) 2000.

7.3. SUMS AND SERIES

THEORY IN A NUTSHELL

In the previous section we talked about the sums of sequences. Often the summed sequence is very long, and sometimes even infinite. Writing in such a situation the sum with the symbol "+" can be confusing. Therefore, we use often the symbol "Σ" (Sigma). For example, the sum of n elements of the sequence (a_n) we write down as follows:

$$a_1 + a_2 + ... + a_n = \sum_{i=1}^{n} a_i .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The right hand side expression we read "the sum of a_i for i equal from 1 to n". "From 1 to n" means that we substitute i with numbers 1, 2, … till n, and the symbol "Σ" represents the addition. The right hand side expression is a short form of the left hand side expression.
Similarly we can write the sum of an infinite sequence:

$$a_1 + a_2 + ... + a_n + ... = \sum_{i=1}^{\infty} a_i \, .$$

The left hand side expression (the sum of the terms of a sequence) is called **series**, and the value of this expression is called the **sum of the series**.
We say that a series is **convergent** if its sum is finite, otherwise the series is **divergent**. An example of convergent series is any geometric series, if the common ratio of the corresponding geometric sequence satisfies the condition $|q| < 1$. Remember that the expression $\sum_{i=1}^{\infty} a_i$ denoted both the series and its sum. On the other hand the sum of the series $\sum_{i=1}^{\infty} a_i$ and the sum of the infinite sequence (a_i) are exactly the same thing.
Sometimes you will also meet product of the elements of a sequence. Also in this case we can use the short notation, using this time the symbol "Π" (Pi) instead of "Σ":

$$a_1 \cdot a_2 \cdot ... \cdot a_n = \prod_{i=1}^{n} a_i \, ,$$

$$a_1 \cdot a_2 \cdot ... \cdot a_n \cdot ... = \prod_{i=1}^{\infty} a_i \, .$$

In order to use these symbols in **WolframAlpha®**, you should use the words "sum" and "product", respectively. For example, in order to write the sum from Example 1 below, you should type:

"sum 2^(i+2), i=1 to 5"

(note that the range is given after the main expression). If you want to write

$$\prod_{i=2}^{\infty} \left(1 - \frac{1}{i^2}\right) \, ,$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

you should type

“product (1-1/i^2), i=2 to infinity”.

Finally, if you want to find short form for a formula, just type it. For example, if you want to solve Example 4 below, you should type:

“1+2^3+3^4+…”.

EXAMPLES

Example 1
Write the following expression, not using the symbol “Σ”:

$$\sum_{i=1}^{5} 2^{i+2} .$$

Solution
The range of i are all integers from 1 to 5, so we substitute i with those numbers:

$$\sum_{i=1}^{5} 2^{i+2} = 2^{1+2} + 2^{2+2} + 2^{3+2} + 2^{4+2} + 2^{5+2} = 2^3 + 2^4 + 2^5 + 2^6 + 2^7 .$$

Example 2
Write the following expression, not using the symbol “Σ”:

$$\sum_{i=1}^{3} \frac{(i+2)^2}{5i} .$$

Solution
This time, we have to substitute i with 1, 2 and 3:

$$\sum_{i=1}^{3} \frac{(i+2)^2}{5i} = \frac{(1+2)^2}{5 \cdot 1} + \frac{(2+2)^2}{5 \cdot 2} + \frac{(3+2)^2}{5 \cdot 3} = \frac{3^2}{5} + \frac{4^2}{10} + \frac{5^2}{15} .$$

Example 3
Write the following expression, using the symbol “Σ”:

$$2 + 2 \cdot 2^2 + 3 \cdot 2^3 + 4 \cdot 2^4 .$$

Solution
In order to solve the problem, we have to find a pattern. Note that $2 = 1 \cdot 2^1$, so we can write the expression as $1 \cdot 2^1 + 2 \cdot 2^2 + 3 \cdot 2^3 + 4 \cdot 2^4$. Both the exponent and the coefficient multiplied by the power change from 1 to 4. Each of them can be thus written as $i \cdot 2^i$, and their sum as:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$2+2\cdot 2^2+3\cdot 2^3+4\cdot 2^4=\sum_{i=1}^{4} i\cdot 2^i \ .$$

Example 4
Write the following expression, using the symbol "Σ":

$$1+2^3+3^4+4^5+5^6+6^7...$$

Solution
Observe that the consecutive integers are raised to a power of exponent greater by one than the base (1 can be written in the form 1^2). In such a case the general form of summand is: i^{i+1}. Note that there are dots in the end. It means that the sequence is infinite. Finally the sum can be written as:

$$1+2^3+3^4+4^5+5^6+6^7...=\sum_{i=1}^{\infty} i^{i+1} \ .$$

Example 5
Write the following expression, using the symbol "Σ":

$$\frac{10}{1{,}1}+\frac{10}{(1{,}1)^2}+\frac{10}{(1{,}1)^3}+\frac{10}{(1{,}1)^4}+\frac{10}{(1{,}1)^5}+\frac{10}{(1{,}1)^6}+\frac{10}{(1{,}1)^7}+\frac{10}{(1{,}1)^8}+\frac{10}{(1{,}1)^9}+\frac{10}{(1{,}1)^{10}} \ .$$

Solution
This time the exponent of the power in denominator changes from 1 to 10. The sum can be written in the form:

$$\frac{10}{1{,}1}+\frac{10}{(1{,}1)^2}+...+\frac{10}{(1{,}1)^{10}}=\sum_{i=1}^{10}\frac{10}{(1{,}1)^i} \ .$$

EXERCISES

1. Write the following expression, not using the symbol "Σ":

   a) $\sum_{i=1}^{5} i^2$;

   b) $\sum_{i=1}^{\infty} i^3$;

   c) $\sum_{k=1}^{\infty}\frac{100}{(1+i)^k}$;

   d) $\sum_{k=1}^{10}\frac{100\cdot 1{,}1^{k-1}}{1{,}2^k}$;

   e) $\sum_{j=1}^{3} j\cdot 2^j$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Write the following expression, using the symbol "Σ":

a) $2+\frac{2}{5}+\frac{2}{25}+\frac{2}{125}+\frac{2}{625}$;

b) $\frac{100}{1{,}1}+\frac{100}{1{,}1^2}+\ldots+\frac{100}{1{,}1^{10}}$;

c) $2+\frac{2}{5}+\frac{2}{25}+\frac{2}{125}+\frac{2}{625}\ldots$;

d) $\frac{100}{1{,}1}+\frac{100}{1{,}1^2}+\frac{100}{1{,}1^3}+\frac{100}{1{,}1^4}+\ldots$;

e) $\frac{100}{1{,}1}+\frac{100\cdot 1{,}05}{1{,}1^2}+\frac{100\cdot 1{,}05^2}{1{,}1^3}+\frac{100\cdot 1{,}05^3}{1{,}1^4}+\ldots$

SOLUTIONS

1. The desired expressions:

a) $1^2+2^2+3^2+4^2+5^2$;

b) $1^3+2^3+3^3+4^3+\ldots$;

c) $\frac{100}{1+i}+\frac{100}{(1+i)^2}+\frac{100}{(1+i)^3}+\ldots$;

d) $\frac{100}{1{,}2}+\frac{100\cdot 1{,}1}{1{,}2^2}+\frac{100\cdot 1{,}1^2}{1{,}2^3}+\frac{100\cdot 1{,}1^3}{1{,}2^4}+\ldots+\frac{100\cdot 1{,}1^8}{1{,}2^9}+\frac{100\cdot 1{,}1^9}{1{,}2^{10}}$;

e) $2+2\cdot 2^2+3\cdot 2^3$.

2. The desired expressions:

a) $\sum_{i=1}^{5}\frac{2}{5^{i-1}}$;

b) $\sum_{i=1}^{10}\frac{100}{1{,}1^i}$;

c) $\sum_{i=1}^{\infty}\frac{2}{5^{i-1}}$;

d) $\sum_{i=1}^{\infty}\frac{100}{1{,}1^i}$;

e) $\sum_{i=1}^{\infty}\frac{100\cdot 1{,}05^{i-1}}{1{,}1^i}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 8
FUNCTIONS OF ONE VARIABLE

8.1. LIMITS

THEORY IN A NUTSHELL

Similarly as in the case of sequences, also in the case of arbitrary functions you can find their limits. In case of a function of real variable these can be, however, not only the limits at ∞ (as in the case of a sequence), but also at –∞ and at a point (i.e., when x approaches some constant). The interpretation of limits at ∞ and –∞ is the same as in the case of limit of a sequence – these are the values to which the value of a function gets arbitrarily close, when x increases (or decreases) indefinitely. In the case of x approaching some number c, we distinguish the following cases:

- **left-sided limit**: the limit when x approaches c from below (left) – the number which the value of the function approaches, when x approaches c and is lower than c;
- **right-sided limit**: the limit when x approaches c from above (right) – the number which the value of the function approaches, when x approaches c and is greater than c;
- **two-sided limit**: the limit when x approaches c – it is equal to two above limits, if they both exist and are the same.

Examples of limits are presented in Fig. 8.1. In the part (a) symbol g denotes the limit at c; in part (b) g_l is the left-sided limit, and g_r – right-sided. Note that $g_l \neq g_r$, so the second function does not have a limit (two-sided) at c.

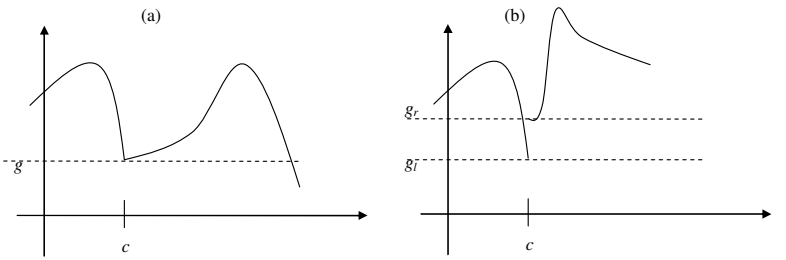


Fig. 8.1. Limits of functions

Instead of saying that a function has the limit equal to ∞ or –∞, it is more proper to say that it **diverges** or **increases/decreases without bound**. An example of

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

divergent function is shown in Fig. 8.2 (the function diverges at c – it approaches ∞ when x approaches c).

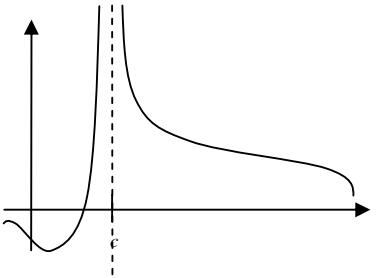


Fig. 8.2. Divergent function

We write the limits of $f(x)$ in similar way as in the case of sequences: limits at ∞ and –∞ as $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ respectively, and two-sided, left-sided and right-sided limit at c as $\lim_{x\to c} f(x)$, $\lim_{x\to c^-} f(x)$ and $\lim_{x\to c^+} f(x)$, respectively.
If you want to find the limit of function $f(x)$ using **WolframAlpha®**, you should use the command "limit". For example, if you want to find the limit

$$\lim_{x\to 4^+} \frac{1}{x-4},$$

you should type:

"limit 1/(x-4) as x-> 4+".

EXAMPLES

Example 1
Find the limit:

$$\lim_{x\to-\infty} x(\sqrt{4x^2+1}-2x).$$

Solution
Let us start with substituting –∞ for x (of course all the "illegal" expressions will be placed in the square brackets). Remember that to the expressions with 0 and ∞ we apply the same rules as in the case of sequences.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim_{x\to-\infty} x(\sqrt{4x^2+1}-2x) = [-\infty(\sqrt{4(-\infty)^2+1}-2(-\infty))] = [-\infty(\sqrt{4\infty+1}+2\infty)] =$$
$$= [-\infty(\infty+\infty)] = [-\infty^2] = -\infty.$$

It works. Unfortunately, it will be not that easy in the next example.

Example 2
Find the limit:

$$\lim_{x\to\infty} x(\sqrt{4x^2+1}-2x).$$

Solution
If we tried the same method as before, we would obtain:

$$\lim_{x\to\infty} x(\sqrt{4x^2+1}-2x) = [\infty(\sqrt{4\infty^2+1}-2\infty)] = [\infty(\infty-\infty)].$$

As you can see, in the end we obtain one of the seven indeterminate forms: $\infty - \infty$ (if you do not remember what are the indeterminate forms, go back to Section 7.1). For this reason, we need to approach this problem differently. How to? Just as in the case of sequences – we multiply and divide function by the conjugate of the expression in parentheses, and then we use the short multiplication formula.

$$\lim_{x\to\infty} x(\sqrt{4x^2+1}-2x) = \lim_{x\to\infty} \frac{x(\sqrt{4x^2+1}-2x)(\sqrt{4x^2+1}+2x)}{(\sqrt{4x^2+1}+2x)} =$$
$$= \lim_{x\to\infty} \frac{x(4x^2+1-4x^2)}{(\sqrt{4x^2+1}+2x)} = \lim_{x\to\infty} \frac{x}{(\sqrt{x^2(4+1/x^2)}+2x)} =$$
$$= \lim_{x\to\infty} \frac{x}{(\sqrt{4x^2}+2x)} = \lim_{x\to\infty} \frac{x}{(2x+2x)} = \frac{1}{4}.$$

Example 3
Find the limit:

$$\lim_{x\to1} \frac{x^3-1}{x^5-x}.$$

Solution
Let us try to substitute (note that this time we are looking for the limit at $x = 1$):

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim_{x\to 1}\frac{x^3-1}{x^5-x}=\left[\frac{1^3-1}{1^5-1}\right]=\left[\frac{0}{0}\right].$$

As a result we obtain the indeterminate form [0/0]. In this situation, before we continue, we need to transform the expression. Since we are dealing with a rational function, we expand the numerator and denominator into prime factors (note that we do not factor out the highest power before the bracket, it is because we now try to find the limit at a point, and not at infinity). For this purpose, we use the short multiplication formulae:

$$\lim_{x\to 1}\frac{x^3-1}{x^5-x}=\lim_{x\to 1}\frac{(x-1)(x^2+x+1)}{x(x^4-1)}=\lim_{x\to 1}\frac{(x-1)(x^2+x+1)}{x(x^2+1)(x^2-1)}=$$

$$=\lim_{x\to 1}\frac{(x-1)(x^2+x+1)}{x(x^2+1)(x+1)(x-1)}=...$$

As you can see, there is a common factor in the numerator and denominator: $(x-1)$. Thanks to that we can simplify the fraction, and try to substitute again:

$$...=\lim_{x\to 1}\frac{(x^2+x+1)}{x(x^2+1)(x+1)}=\frac{(1^2+1+1)}{1(1^2+1)(1+1)}=\frac{3}{4}.$$

Example 4
Find the limit:

$$\lim_{x\to 2}\frac{\sqrt{20x^2+1}-9}{x-2}.$$

Solution
Again, as in the previous example, after substituting we obtain an indeterminate form:

$$\lim_{x\to 2}\frac{\sqrt{20x^2+1}-9}{x-2}=\left[\frac{\sqrt{20\cdot 2^2+1}-9}{2-2}\right]=\left[\frac{0}{0}\right].$$

In order to find the limit we have to transform the formula. How to? Just as in the previous cases when we dealt with subtracting the roots – we multiply and divide by the conjugate:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim_{x\to 2}\frac{\sqrt{20x^2+1}-9}{x-2}=\lim_{x\to 2}\left(\frac{\sqrt{20x^2+1}-9}{x-2}\cdot\frac{\sqrt{20x^2+1}+9}{\sqrt{20x^2+1}+9}\right)=$$

$$=\lim_{x\to 2}\frac{20x^2+1-81}{(x-2)(\sqrt{20x^2+1}+9)}=\lim_{x\to 2}\frac{20(x^2-4)}{(x-2)(\sqrt{20x^2+1}+9)}=$$

$$=\lim_{x\to 2}\frac{20(x+2)(x-2)}{(x-2)(\sqrt{20x^2+1}+9)}=...$$

As you can see, there is a common factor $(x-2)$ and we can simplify the expression. Then we substitute 2 for x and we obtain:

$$...=\lim_{x\to 2}\frac{20(x+2)}{\sqrt{20x^2+1}+9}=\frac{20(2+2)}{\sqrt{20\cdot 2^2+1}+9}=\frac{40}{9}.$$

Example 5
Find the limit:

$$\lim_{x\to 0}(1-7x)^{\frac{3}{x}}.$$

Solution
If we substitute 0 for x, then we obtain again an indeterminate form: $[1^{\infty}]$. Recall from Section 7.1 that in such a case the expression or its part approaches e. The only thing that we need is the exponent being the reciprocal of the expression added to 1 and appropriately chosen coefficients:

$$\lim_{x\to 0}(1-7x)^{\frac{3}{x}}=\lim_{x\to 0}(1-7x)^{\frac{1}{-7x}\cdot(-21)}=\lim_{x\to 0}\left[(1-7x)^{\frac{1}{-7x}}\right]^{-21}=e^{-21}.$$

Example 6
Find the limit:

$$\lim_{x\to 0}\ln(x^4).$$

Solution
This time it is enough to know the properties of the function ln x (if you do not remember them, go back to Section 6.4):

$$\lim_{x\to 0}\ln(x^4)=[\ln 0^+]=-\infty.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that we wrote "0^+" instead of "0" in the square brackets. We can do this since $x^4 \geq 0$ for each $x \in R$. On the other hand it is necessary, because the function $\ln x$ is defined only for positive x. Recall that "0^+" means that given expression approaches 0 from above, i.e. it approaches 0 and is greater than 0.

Example 7
Find the limit:

$$\lim_{x \to 4^+} \frac{1}{x-4}.$$

Solution
We will use the basic rules for the operations on 0 and ∞. Namely we will use the property $[1/0^+] = \infty$:

$$\lim_{x \to 4^+} \frac{1}{x-4} = \left[\frac{1}{4^+ - 4}\right] = \left[\frac{1}{0^+}\right] = \infty .$$

Example 8
Find the limit:

$$\lim_{x \to 4^-} \frac{1}{x-4}.$$

Solution
We will solve this example just like the previous one, this time using the property $[1/0^-] = -\infty$:

$$\lim_{x \to 4^-} \frac{1}{x-4} = \left[\frac{1}{4^- - 4}\right] = \left[\frac{1}{0^-}\right] = -\infty .$$

Example 9
Find the limit:

$$\lim_{x \to 0^+} \frac{e^{\frac{1}{x}} - 4}{e^{\frac{1}{x}} + 4}.$$

Solution
We will use the rules: $[1/0^+] = \infty$ and $[a^\infty] = \infty$, for $a > 1$:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\lim_{x\to 0^+}\frac{e^{\frac{1}{x}}-4}{e^{\frac{1}{x}}+4}=\left[\frac{e^{\frac{1}{0^+}}-4}{e^{\frac{1}{0^+}}+4}\right]=\left[\frac{e^{\infty}-4}{e^{\infty}+4}\right]=\left[\frac{\infty}{\infty}\right].$$

Note that we obtained the indeterminate form [∞/∞]. Since we are dealing with a fraction, in the numerator and the denominator of which there are powers (which tend to ∞), we factor out the highest power and obtain the result:

$$\lim_{x\to 0^+}\frac{e^{\frac{1}{x}}-4}{e^{\frac{1}{x}}+4}=\lim_{x\to 0^+}\frac{e^{1/x}(1-4/e^{1/x})}{e^{1/x}\left(1+4/e^{1/x}\right)}=\lim_{x\to 0^+}\frac{1-4/e^{1/x}}{1+4/e^{1/x}}=\left[\frac{1-4/\infty}{1+4/\infty}\right]=\frac{1-0}{1+0}=1.$$

Example 10
Find the limit:

$$\lim_{x\to 0^-}\frac{e^{\frac{1}{x}}-4}{e^{\frac{1}{x}}+4}.$$

Solution
We will begin as in the previous example. This time we will use the property of powers $a^{-x} = 1/a^x$.

$$\lim_{x\to 0^-}\frac{e^{\frac{1}{x}}-4}{e^{\frac{1}{x}}+4}=\left[\frac{e^{\frac{1}{0^-}}-4}{e^{\frac{1}{0^-}}+4}\right]=\left[\frac{e^{-\infty}-4}{e^{-\infty}+4}\right]=\left[\frac{1/e^{\infty}-4}{1/e^{\infty}+4}\right]=\frac{0-4}{0+4}=-1.$$

As you can see, this time we managed to get the result without additional transformations.

EXERCISES

1. Find the limits:
   a) $\lim_{x\to\infty} x(\sqrt{2x^2+2x}-\sqrt{2x^2+x})$;
   b) $\lim_{x\to-\infty}(\sqrt{x^2-2x}-\sqrt{x^2-x})$;
   c) $\lim_{x\to-7}\frac{x^2-49}{x+7}$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

d) $\lim\limits_{x\to 2}\dfrac{x^4-16}{x^3-8}$;

e) $\lim\limits_{x\to 4}\dfrac{\sqrt{x^2+9}-\sqrt{6x+1}}{x^2-16}$;

f) $\lim\limits_{x\to 2}\dfrac{\sqrt{x^2+5}-\sqrt{6x-3}}{\sqrt{6x^2+1}-\sqrt{9x^2-11}}$;

g) $\lim\limits_{x\to 0}(1+4x)^{\frac{2}{x}}$;

h) $\lim\limits_{x\to 0}\left(\dfrac{1+4x}{1+5x}\right)^{\frac{3}{x}}$;

i) $\lim\limits_{x\to -\infty}\dfrac{3+4^x}{2+5^x}$;

j) $\lim\limits_{x\to \infty}\dfrac{3+4^x}{2+5^x}$;

k) $\lim\limits_{x\to \infty}\ln\left(\dfrac{2x}{x^2+1}\right)$;

l) $\lim\limits_{x\to \infty}\ln\left(\dfrac{2x^3}{x^2+1}\right)$.

2. Find the one-sided limits of the functions:

a) $f(x)=\dfrac{1}{x-10}$ at $x = 10$;

b) $f(x)=\dfrac{1}{x^2-64}$ at $x = -8$;

c) $f(x)=\dfrac{3^{\frac{1}{x}}-3}{3^{\frac{1}{x}}+3}$ at $x = 0$;

d) $f(x)=7^{\frac{5}{x^2-9}}$ at $x = 3$;

e) $f(x)=\ln(2^x)$ at $x = 0$.

SOLUTIONS

1. a) ∞; b) 1/2; c) –14; d) 8/3;
   e) 1/40; f) 5/18; g) e^8; h) e^{-3};
   i) 3/2; j) 0; k) –∞; l) ∞.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. a) $\lim\limits_{x\to 10^-}\dfrac{1}{x-10}=-\infty$, $\lim\limits_{x\to 10^+}\dfrac{1}{x-10}=\infty$;

b) $\lim\limits_{x\to -8^-}\dfrac{1}{x^2-64}=\infty$, $\lim\limits_{x\to -8^+}\dfrac{1}{x^2-64}=-\infty$;

c) $\lim\limits_{x\to 0^-}\dfrac{3^{\frac{1}{x}}-3}{3^{\frac{1}{x}}+3}=-1$, $\lim\limits_{x\to 0^+}\dfrac{3^{\frac{1}{x}}-3}{3^{\frac{1}{x}}+3}=1$;

d) $\lim\limits_{x\to 3^-}7^{\frac{5}{x^2-9}}=0$, $\lim\limits_{x\to 3^+}7^{\frac{5}{x^2-9}}=\infty$;

e) $\lim\limits_{x\to 0^-}\ln\left(2^x\right)=0$, $\lim\limits_{x\to 0^+}\ln\left(2^x\right)=0$.

8.2. DERIVATIVES

THEORY IN A NUTSHELL

One of the most important notions in the economic analysis are increments, measured by the **derivatives of the functions**. The derivative of function $f(x)$ is denoted by $f'(x)$. Sometimes it is better to use another notation: $\dfrac{dy}{dx}$, when we are talking about the function $y=f(x)$. The derivative at each point takes the value equal to the slope of the graph of the function[20]. More formally, the derivative is defined as follows.

Definition 8.1. Derivative
The derivative of $f(x)$ at x_0, if it exists, is equal to the following limit of the **difference quotient**:

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}.$$

Fortunately, you do not need to calculate the derivatives using the above definition – it has been already done for you. In Table 8.1 the formulae for the most important functions are given.

[20] To be more specific, it is equal to the slope of the tangent at this point.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Table 8.1. Derivatives (a, c – constants, e – number e)

Function $f(x)$	Derivative $f'(x)$	Function $f(x)$	Derivative $f'(x)$
c	0	$\ln x$	$1/x$
x	1	e^x	e^x
x^a	ax^{a-1}	a^x	$a^x \cdot \ln a$

Of course, this is not a complete list of functions, which we will deal with. To determine the derivatives of more complex ones, you can use the following properties.

Theorem 8.1. Rules of computation

1. Sum rule: $(af(x)+bg(x))' = af'(x)+bg'(x)$.
2. Product rule: $(f(x)g(x))' = f'(x)g(x)+f(x)g'(x)$.
3. Quotient rule: $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$.
4. Chain rule: $[f(g(x))]' = f'(g(x))g'(x)$.

One of the most important applications of derivatives is to determine the **extrema** of a function and study its **monotonicity**. We will use the following properties.

Theorem 8.2. Derivatives, extrema and monotonicity

1. The derivative is positive on an interval (a, b) if and only if the function is increasing on this interval.
2. The derivative is negative on an interval (a, b) if and only if the function is decreasing on this interval.
3. If at some $x \in D(f)$ the function f has an extremum, then either $f'(x) = 0$ (**type 1 stationary point**), or $x \notin D(f')$ (**type 2 stationary point**).

Given the above relations, the procedure for determining the extrema and monotonicity test is as follows:

1. Find the domains of the function and its derivatives (note: $D(f') \subset D(f)$ – if it does not follow directly from the formulae, you have to restrict $D(f')$). If $D(f) \setminus D(f') \neq \emptyset$, then all $x \in D(f) \setminus D(f')$ are the type 2 stationary points.
2. Solve the equation $f'(x) = 0$ (remember to take into account the domain $D(f')$). All its solutions are the type 2 stationary points.
3. Solve the inequality $f'(x) > 0$ (remember to take into account the domain $D(f')$). The function increases on each of the intervals, which the solution of the inequality consists of.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

4. Solve the inequality $f'(x) < 0$ (remember to take into account the domain $D(f')$). The function decreases on each of the intervals, which the solution of the inequality consists of.
5. Examine all the stationary points (type 1 and 2). Check whether the function increases or decreases to the right and left of each such point and determine the version of the extremum, if it exists.

Second derivative of a function f is the derivative of its derivative. We denote it by $f''(x)$. The second derivatives also can be used to verify the existence of extremum of the function: if the value of second derivative at a stationary point is positive, the function has local minimum at this point, and if it is negative, the function has a maximum. In this chapter, we will not use this method of verification, as it often does not give a clear answer, and in some cases (for example in the case of the type 2 stationary points), it cannot be used. The second derivative will be useful to us, however, in Chapter 9.

If you want to find the derivative of function $f(x)$ using **WolframAlpha®**, you can use the command "derivative" (together with the word "second", if necessary). In order to find the extrema, you should use the command "extrema". When you find the extrema, you can read the monotonicity from the plot, which is usually attached to the report. For example, if you want to find the derivative of the function from Example 1, you should type:

"derivative 4x^3-2x+17",

while in order to find the second derivative of the same function you could use the command:

"second derivative 4x^3-2x+17".

In order to find the extrema of this function, you can use:

"extrema 4x^3-2x+17".

EXAMPLES

Example 1
Find the derivative of the function:

$f(x) = 4x^3 - 2x + 17$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
We will use the formulae for the derivatives of a power function, linear function and constant: $(4x^3)' = 4 \cdot 3x^2$, $(-2x)' = -2 \cdot 1$, $(17)' = 0$, thus finally:

$$f'(x) = 12x^2 - 2 .$$

Example 2
Find the derivative of the function:

$$f(x) = e^{x^2} .$$

Solution
The derivative of e^x is e^x. However, there is x^2 in the exponent, instead of x. Thus we deal with a composite function and we have to apply the chain rule. The derivative of $f(x)$ is thus equal to derivative of the outer function e^{x^2} multiplied by the derivative of the inner function x^2, i.e. $2x$:

$$f'(x) = e^{x^2} \cdot 2x .$$

Example 3
Find the derivative of the function:

$$f(x) = (2x^2 - x)e^{-x} .$$

Solution
This time the function has the form of the product of two functions: $2x^2 - x$ and e^{-x}. We apply the product rule:

$$f'(x) = (2x^2 - x)'e^{-x} + (2x^2 - x)(e^{-x})' .$$

In order to find the derivative of the first factor, we apply the formulae for the derivatives of the power and linear functions:

$$(2x^2 - x)' = 2 \cdot 2x - 1 .$$

In the second case we apply the formula for the derivative of e^x, bearing in mind that it is necessary to use the chain rule as the inner function is $-x$:

$$(e^{-x})' = e^{-x} \cdot (-1) .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Finally the derivative of $f(x)$ has the form:

$$f'(x) = (4x-1)e^{-x} + (2x^2 - x)(-e^{-x}) = (-2x^2 + 5x - 1)e^{-x}.$$

Example 4
Find the derivative of the function:

$$f(x) = \frac{\ln x^2}{e^x}.$$

Solution
The function is given as a quotient, so we apply the quotient rule:

$$f'(x) = \frac{(\ln x^2)'e^x - (\ln x^2)(e^x)'}{(e^x)^2}.$$

The derivative of e^x is e^x. The derivative of $\ln x$ is in turn $1/x$. By applying the chain rule (the inner function is x^2), we obtain:

$$(\ln x^2)' = \frac{1}{x^2} \cdot 2x.$$

Thus finally we obtain:

$$f'(x) = \frac{\frac{1}{x^2} \cdot 2x \cdot e^x - (\ln x^2)e^x}{(e^x)^2} = \frac{\frac{2}{x} - \ln x^2}{e^x} = \frac{2 - x\ln x^2}{xe^x}.$$

Example 5
Find the derivative of the function:

$$f(x) = \sqrt{\frac{16 - x^2}{16 + x^2}}.$$

Solution
We begin with writing the root in the form of a power:

$$f(x) = \left(\frac{16 - x^2}{16 + x^2}\right)^{1/2}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Now we apply the formula for the derivative of a power, bearing in mind that there is an inner function inside the brackets:

$$f'(x) = \frac{1}{2}\left(\frac{16-x^2}{16+x^2}\right)^{-1/2} \cdot \left(\frac{16-x^2}{16+x^2}\right)'.$$

The inner function has the form of a quotient, so we apply the quotient rule:

$$\left(\frac{16-x^2}{16+x^2}\right)' = \frac{(16-x^2)'(16+x^2)-(16-x^2)(16+x^2)'}{(16+x^2)^2} =$$

$$= \frac{-2x(16+x^2)-(16-x^2)\cdot 2x}{(16+x^2)^2}.$$

Finally we obtain:

$$f'(x) = \frac{1}{2}\left(\frac{16-x^2}{16+x^2}\right)^{-1/2} \cdot \frac{-2x(16+x^2)-(16-x^2)\cdot 2x}{(16+x^2)^2} = \frac{-32x}{(16-x^2)^{1/2}(16+x^2)^{3/2}}$$

.

Example 6
Determine the intervals of increase and decrease and find all the local extrema of the function:

$$f(x) = x^3 - 6x^2 + 9x + 25.$$

Solution
The domain is the set of the real numbers: $D(f) = R$. The derivative has the form:

$$f'(x) = 3x^2 - 12x + 9.$$

The domain of the derivative is the same as the domain of the function: $D(f') = R$, so there are no type 2 stationary points. In order to find the type 1 stationary points we must solve the equation[21] $f'(x) = 0$:

$$3x^2 - 12x + 9 = 0.$$

[21] If you do not remember how to solve the equations and inequalities like this, read again Chapter 6.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

It follows that $x = 1 \vee x = 3$. Thus the function has two type 1 stationary points.
In order to determine its monotonicity, we solve the inequalities: $f'(x) > 0$ (i.e. $3x^2 - 12x + 9 > 0$) and $f'(x) < 0$ (i.e. $3x^2 - 12x + 9 < 0$). The solution of the first one is $(-\infty, 1) \cup (3, \infty)$, the solution of the second one is $(1, 3)$.
We will summarize the results in Table 8.2. In the first row we list all the important values of x (here: stationary points, in the other examples these can be also the ends of the domain) and intervals corresponding with the remaining values of x. In the second row we schematically indicate where the derivative is positive ("+"), negative ("–") and equal to 0 ("0"). If it is not defined at some point, then we use the sign "×".

Table 8.2. Properties of the function

x	$(-\infty, 1)$	1	$(1, 3)$	3	$(3, \infty)$
$f'(x)$	+	0	–	0	+
$f(x)$	↑	max	↓	min	↑

The last row corresponds with the function itself. It increases on the intervals where the derivative is positive (we denote it by "↑"), and decreases on the intervals, where the derivative is negative (symbol "↓"). Finally we determine where are the extrema (for example at $x = 1$ there is a local maximum, as the function increases on the left and decreases on the right of this point). We end with calculating the value of the function at all the extrema: $f(1) = 1^3 - 6 \cdot 1^2 + 9 \cdot 1 + 25 = 29$, $f(3) = 3^3 - 6 \cdot 3^2 + 9 \cdot 3 + 25 = 25$. To summarize: the function increases for $x \in (-\infty, 1)$ and for $x \in (3, \infty)$, decreases for $x \in (1, 3)$, it has a local maximum at $x = 1$, $f_{\max}(1) = 29$, and a local minimum at $x = 3$, $f_{\min}(3) = 25$.
Remark: note that we do not write that the function increases for $x \in (-\infty, 1) \cup (3, \infty)$, but that it increases for $x \in (-\infty, 1)$ and for $x \in (3, \infty)$.

Example 7
Determine the intervals of increase and decrease and find all the local extrema of the function:

$$f(x) = \sqrt{16 - x^2}\,.$$

Solution
The domain is the interval $[-4, 4]$. We calculate the derivative:

$$f(x)' = \left[\left(16 - x^2\right)^{1/2}\right]' = \frac{1}{2}\left(16 - x^2\right)^{-1/2} \cdot (-2x) = \frac{-x}{\sqrt{16 - x^2}}\,.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The domain of the derivative is $(-4, 4)$. Note that $D(f) \setminus D(f') = \{-4, 4\}$, so the $x = -4$ and $x = 4$ are type 2 stationary points.
By equating the derivative to 0, we obtain:

$$\frac{-x}{\sqrt{16-x^2}} = 0 \quad \Leftrightarrow \quad -x = 0 \quad \Leftrightarrow \quad x = 0\,.$$

Hence $x = 0$ is a type 1 stationary point. In order to determine the monotonicity, we solve the inequalities (remember that $x \in (-4, 4)$):

$$f'(x) > 0 \quad \Leftrightarrow \quad \frac{-x}{\sqrt{16-x^2}} > 0 \quad \Leftrightarrow \quad x \in (-4, 0)\,,$$

$$f'(x) < 0 \quad \Leftrightarrow \quad \frac{-x}{\sqrt{16-x^2}} < 0 \quad \Leftrightarrow \quad x \in (0, 4)\,.$$

The results are presented in Table 8.3. Observe that the intervals $(-\infty, -4)$ and $(4, \infty)$, on which the function is not defined, have been skipped. Note also that the function has extrema at the ends of the domain.

Table 8.3. Properties of the function

x	-4	$(-4, 0)$	0	$(0, 4)$	4
$f'(x)$	×	+	0	–	×
$f(x)$	min	↑	max	↓	min

To summarize, $f(x)$ increases for $x \in (-4, 0)$, decreases for $x \in (0, 4)$, has a local minimum at $x = -4$, $f_{\min}(-4) = 0$, a local minimum at $x = 4$, $f_{\min}(4) = 0$ and a local maximum at $x = 0$, $f_{\max}(0) = 4$.

Example 8
Determine the intervals of increase and decrease and find all the local extrema of the function:

$$f(x) = |x^2 - 8x + 12|\,.$$

Solution
Since the function has been defined with the absolute value, we have to consider two cases: when the expression inside is nonnegative (when $x^2 - 8x + 12 \geq 0$, i.e. when $x \in (-\infty, 2] \cup [6, \infty)$), and when it is negative (when $x^2 - 8x + 12 < 0$, i.e. when $x \in (2, 6)$). Taking into account the definition of the absolute value:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$|a| = \begin{cases} a, & \text{if } a \geq 0, \\ -a, & \text{if } a < 0, \end{cases}$$

we can write the function as:

$$f(x) = \begin{cases} x^2 - 8x + 12, & \text{if } x \in (-\infty, 2] \cup [6, \infty), \\ -(x^2 - 8x + 12), & \text{if } x \in (2, 6). \end{cases}$$

The domain is the set of real numbers: $D(f) = R$. We find the derivative for both cases separately:

$$f'(x) = \begin{cases} 2x - 8, & \text{if } x \in (-\infty, 2) \cup (6, \infty), \\ -2x + 8, & \text{if } x \in (2, 6). \end{cases}$$

Note that the intervals are open now. The reason is that the derivative is always defined on open intervals. It follows that $x = 2$ and $x = 6$ are type 2 stationary points. We perform next steps separately for both cases.
Case 1: $x \in (-\infty, 2) \cup (6, \infty)$. The derivative equals to $f'(x) = 2x - 8$. Taking into account the assumption that $x \in (-\infty, 2) \cup (6, \infty)$, we obtain the solutions: $f'(x) = 0$ – no solution, $f'(x) > 0$ for $x \in (6, \infty)$, $f'(x) < 0$ for $x \in (-\infty, 2)$.
Case 2: $x \in (2, 6)$. The derivative equals to $f'(x) = -2x + 8$. Taking into account the assumption that $x \in (2, 6)$, we obtain the solutions: $f'(x) = 0$ at $x = 4$, $f'(x) > 0$ for $x \in (2, 4)$, $f'(x) < 0$ for $x \in (4, 6)$.
The summary has been presented in Table 8.4.

Table 8.4. Properties of the function

x	$(-\infty, 2)$	2	$(2, 4)$	4	$(4, 6)$	6	$(6, \infty)$
$f'(x)$	–	×	+	0	–	×	+
$f(x)$	↓	min	↑	max	↓	min	↑

As we can see, the function increases for $x \in (2, 4)$ and for $x \in (6, \infty)$, it decreases for $x \in (-\infty, 2)$ and for $x \in (4, 6)$, it has a local minimum at $x = 2$, $f_{\min}(2) = 0$, a local minimum at $x = 6$, $f_{\min}(6) = 0$, and a local maximum at $x = 4$, $f_{\max}(4) = 4$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXERCISES

1. Find the derivative of the function:
   a) $f(x) = 5x^4 + 3x^2 + 17$;
   b) $f(x) = \dfrac{x^3 - 2x}{2x + 2}$;
   c) $f(x) = \dfrac{x^2 - 8}{x^2 + 8}$;
   d) $f(x) = \sqrt{\dfrac{x+5}{x-5}}$;
   e) $f(x) = \sqrt{\dfrac{x^2+5}{x^2-5}}$;
   f) $f(x) = (x^2 - 1)\ln x$;
   g) $f(x) = \ln(2x^2 - 8)$;
   h) $f(x) = e^{x^2-10}$;
   i) $f(x) = e^x(x^4 + 16)$;
   j) $f(x) = \dfrac{e^{2x}}{x^5}$.
2. Determine the intervals of increase and decrease and find all the local extrema of the function:
   a) $f(x) = x^2 + 8x + 6$;
   b) $f(x) = 2x^3 - 9x^2 - 24x + 100$;
   c) $f(x) = \dfrac{5}{x^2 - 25}$;
   d) $f(x) = \dfrac{2x+3}{x^2+4}$;
   e) $f(x) = (x+7)e^{x+7}$;
   f) $f(x) = e^{-10x^2}$;
   g) $f(x) = \ln(x^2 - 10x)$;
   h) $f(x) = x^2 \ln x$;
   i) $f(x) = \left|2x^2 - 10x\right|$;
   j) $f(x) = \sqrt[3]{(x-5)^2}$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. a) $20x^3 + 6x$; b) $\dfrac{4x^3 + 6x^2 - 4}{(2x+2)^2}$;

   c) $\dfrac{32x}{\left(x^2+8\right)^2}$; d) $\dfrac{-5}{(x-5)^{3/2}(x+5)^{1/2}}$;

   e) $\dfrac{-10x}{\left(x^2-5\right)^{3/2}\left(x^2+5\right)^{1/2}}$; f) $2x\ln x + \dfrac{x^2-1}{x}$;

   g) $\dfrac{2x}{x^2-4}$; h) $2xe^{x^2-10}$;

   i) $e^x(x^4+4x^3+16)$; j) $\dfrac{e^{2x}(2x-5)}{x^6}$.

2. Summary:
   a) increases for $x \in (-4, \infty)$, decreases for $x \in (-\infty, -4)$, $f_{\min}(-4) = -10$;
   b) increases for $x \in (-\infty, -1)$ and for $x \in (4, \infty)$, decreases for $x \in (-1, 4)$, $f_{\max}(-1) = 100$, $f_{\min}(4) = -12$;
   c) increases for $x \in (-\infty, -5)$ and for $x \in (-5, 0)$, decreases for $x \in (0, 5)$ and for $x \in (5, \infty)$, $f_{\max}(0) = -1/5$;
   d) increases for $x \in (-4, 1)$, decreases for $x \in (-\infty, -4)$ and for $x \in (1, \infty)$, $f_{\min}(-4) = -1/4$, $f_{\max}(1) = 1$;
   e) increases for $x \in (-8, \infty)$, decreases for $x \in (-\infty, -8)$, $f_{\min}(-8) = -1/e$;
   f) increases for $x \in (-\infty, 0)$, decreases for $x \in (0, \infty)$, $f_{\max}(0) = 1$;
   g) increases for $x \in (10, \infty)$, decreases for $x \in (-\infty, 0)$, no extrema;
   h) increases for $x \in (e^{-1/2}, \infty)$, decreases for $x \in (0, e^{-1/2})$, $f_{\min}(e^{-1/2}) = -1/(2e)$;
   i) increases for $x \in (0, 5/2)$ and for $x \in (5, \infty)$, decreases for $x \in (-\infty, 0)$ and for $x \in (5/2, 5)$, $f_{\min}(0) = 0$, $f_{\max}(5/2) = 25/2$, $f_{\min}(5) = 0$;
   j) increases for $x \in (5, \infty)$, decreases for $x \in (-\infty, 5)$, $f_{\min}(5) = 0$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 9
FUNCTIONS OF MANY VARIABLES

9.1. FUNCTIONS OF MANY VARIABLES AND THEIR DERIVATIVES

THEORY IN A NUTSHELL

Function of many variables $f(x_1, x_2, ..., x_n)$ is a function having more than one argument[22]. To determine the value of a function of many variables, you have to know the values of all its arguments, not just one, like so far. For example, consider a function of two variables:

$$f(x_1, x_2) = e^{2x_1}(x_1 + x_2^2 - 4x_2).$$

In order to know its value at some point, it is necessary to know the values of both x_1 and x_2. Let us calculate its value at two points, (0, 1) and (1, 2):

$$f(0,1) = e^{2\cdot 0}(0 + 1^2 + 2\cdot 1) = 3,$$

$$f(1,2) = e^{2\cdot 1}(1 + 2^2 + 2\cdot 2) = 9e^2.$$

Because of the possible applications, we are going to focus only on the differential calculus of this kind of functions. In this case we cannot say "derivative", it is necessary to calculate the **partial derivatives**, i.e. the **derivatives with respect to one of the variables, with the others held constant**. Note that the number of partial derivatives is equal to the number of variables.

Definition 9.1. Partial derivative
Partial derivative of a function f with respect to variable x_j is a derivative calculated under assumption that all the variables except x_j are constant. The partial derivative of f with respect to x_j is denoted by

$$\left.\frac{\partial f}{\partial x_j}\right|(x_1, x_2, ..., x_n)$$

(we usually omit the arguments in this case) or by

[22] We will often use the letters x, y, z, … instead of x with index: x_1, x_2, x_3 …

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$f'_{x_j}(x_1, x_2, ..., x_n)$.

The partial derivatives are elements of the vector called **gradient**.

Definition 9.2. Gradient
Gradient $\nabla f(x_1, x_2, ..., x_n)$ of function *f* is the vector consisting of all its partial derivatives. It is the direction of the fastest increase of *f*.

Just as in the case of functions of one variable, we can define the second order derivatives. Since we can find the derivative of each derivative with respect to each variable, their number equals to the square of the number of variables.

Definition 9.3. Second order derivatives
Second order partial derivative of *f* with respect to x_j is the derivative with respect to x_j of the derivative of *f* with respect to x_j. **Second order mixed derivative** of *f* with respect to x_i and x_j is the derivative with respect to x_i of the derivative of *f* with respect to x_j. We denote them respectively by:

$$\left.\frac{\partial^2 f}{\partial x_j^2}\right|_{(x_1, x_2, ..., x_n)} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right)\Bigg|_{(x_1, x_2, ..., x_n)}$$

and

$$\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{(x_1, x_2, ..., x_n)} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)\Bigg|_{(x_1, x_2, ..., x_n)}$$

(in this case we usually skip the arguments) or by

$f''_{x_j x_j}(x_1, x_2, ..., x_n)$ and $f''_{x_i x_j}(x_1, x_2, ..., x_n)$.

Definition 9.4. Hessian matrix
Hessian matrix (or **Hessian**) $Hf(x_1, x_2, ..., x_n)$ is the square matrix whose elements are the second order partial and mixed derivatives of $f(x_1, x_2, ..., x_n)$. For every pair (*i*, *j*), the element h_{ij} is equal to $f''_{x_i x_j}(x_1, x_2, ..., x_n)$.

Second order mixed derivatives have a useful property, thanks to which it is enough to compute half of them.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Theorem 9.1. Schwarz' Theorem
If at some point both mixed derivatives with respect to x_i and x_j exist, then they are equal:

$$f''_{x_i x_j}(x_1, x_2, ..., x_n) = f''_{x_j x_i}(x_1, x_2, ..., x_n).$$

It follows from the Schwarz' Theorem that Hessian is a symmetric matrix.
In order to implement the derivatives in **WolframAlpha®**, you can use the notation "d/dx", for example to find the derivatives with respect to x and y of the function from Example 2 below, you should type:

"d/dx (exp(y)(8x+x^2-4y)), d/dy (exp(y)(8x+x^2-4y))".

If you use indexed variables (like in Example 1 below), it is necessary to use the word "derivative". For example, you can find the derivative with respect to x_1 by typing:

"derivative x1^2+3x1x2+11x2^2-5x1-25x2 with respect to x1".

EXAMPLES

Example 1
Find the gradient and Hessian of the function:

$$f(x_1, x_2) = x_1^2 + 3x_1x_2 + 11x_2^2 - 5x_1 - 25x_2.$$

Solution
Let us begin with finding the first order derivatives. In order to compute the derivative with respect to x_1, we assume that x_2 is constant. Derivative of the first addend is $2x_1$ (x_2 does not occur here). The second addend is $3x_1x_2$, or $(3x_2)x_1$. As the derivative of cx is c, and this addend is of this form (constant times x), its derivative is $3x_2$. The third addend is a constant, so its derivative is 0. The fourth one has derivative equal to –5, and the last one is again a constant, so its derivative is 0. Finally, the first order partial derivative of f with respect to x_1 is: $f'_{x_1}(x_1, x_2) = 2x_1 + 3x_2 - 5$. By repeating this reasoning for x_2 (now we treat x_1 as constant), we obtain $f'_{x_2}(x_1, x_2) = 3x_1 + 22x_2 - 25$. Taking into account the above, the gradient of f is:

$$\nabla f(x_1, x_2) = [2x_1 + 3x_2 - 5, \quad 3x_1 + 22x_2 - 25]^T.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Note that we use the transposition. It follows from the fact that it is a column vector, but for the sake of convenience it is better to write it as a row.
Nowe we proceed to the second order derivatives (partial and mixed). Let us begin with the derivatives of $f'_{x_1}(x_1, x_2)$. The derivative with respect to x_1 (taking into account that we treat x_2 as constant) equals to:

$$f''_{x_1x_1}(x_1, x_2) = 2 + 0 - 0 = 2.$$

Similarly, the derivative with respect to x_2 equals to:

$$f''_{x_2x_1}(x_1, x_2) = 0 + 3 - 0 = 3.$$

Analogously we find the derivatives of $f'_{x_2}(x_1, x_2)$ – with respect to x_1 and x_2:

$$f''_{x_1x_2}(x_1, x_2) = 3 + 0 - 0 = 3,$$
$$f''_{x_2x_2}(x_1, x_2) = 0 + 22 - 0 = 22.$$

Finally, the Hessian matrix of f is:

$$Hf(x_1, x_2) = \begin{bmatrix} 2 & 3 \\ 3 & 22 \end{bmatrix}.$$

Note that both mixed derivatives are equal to 3 – it is consistent with the Schwarz' Theorem.

Example 2
Find the gradient and Hessian of the function:

$$f(x, y) = e^y(8x + x^2 - 4y).$$

Solution
Let us begin with the derivative with respect to x. Since y is treated as constant, also e^y is constant, so we can focus on finding the derivative of the expression in brackets. It is equal to $8 + 2x$, since we skip the last addend ($4y$ is constant). It follows that the derivative with respect to x equals to:

$$f'_x(x, y) = e^y(8 + 2x).$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In the case of y the first factor (e^y) cannot be treated as constant, so we have to apply the product rule. The derivative of the first factor equals to e^y, and the derivative of the second one is –4, so finally the derivative equals to:

$$f'_y(x,y) = e^y(8x + x^2 - 4y) + e^y \cdot (-4) = e^y(8x + x^2 - 4y - 4).$$

The gradient is:

$$\nabla f(x,y) = [e^y(8+2x), \quad e^y(8x + x^2 - 4y - 4)]^T.$$

Proceeding as before, we determine the derivatives of derivatives, i.e. the second order derivatives. The derivatives of $f'_x(x,y)$ are equal to:

$$f''_{xx}(x,y) = 2e^y,$$

$$f''_{yx}(x,y) = e^y(8+2x),$$

and the derivatives of $f'_y(x,y)$ are:

$$f''_{xy}(x,y) = e^y(8+2x),$$

$$f''_{yy}(x,y) = e^y(8x + x^2 - 4y - 8).$$

Finally, the Hessian is:

$$Hf(x,y) = \begin{bmatrix} 2e^y & e^y(8+2x) \\ e^y(8+2x) & e^y(8x + x^2 - 4y - 8) \end{bmatrix}.$$

Example 3
Find the gradient and Hessian of the function:

$$f(x,y,z) = \ln(x^2 + 4y^2 + 9z^2 + 1).$$

Solution
This time, no matter which variable we choose, we have to deal with a composite function: the outer is logarithmic function (in each case the derivative equals to $\frac{1}{x^2 + 4y^2 + 9z^2 + 1}$), and the inner function is $x^2 + 4y^2 + 9z^2 + 1$, whose derivative depends on the choice of variable: the

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

derivative with respect to x is $2x$, the derivative with respect to y is $8y$, and the derivative with respect to z is $18z$. Finally, the first order partial derivatives are:

$$f'_x(x,y,z)=\frac{2x}{x^2+4y^2+9z^2+1},$$

$$f'_y(x,y,z)=\frac{8y}{x^2+4y^2+9z^2+1},$$

$$f'_z(x,y,z)=\frac{18z}{x^2+4y^2+9z^2+1},$$

and the gradient is:

$$\nabla f(x,y,z)=\left[\frac{2x}{x^2+4y^2+9z^2+1},\quad \frac{8y}{x^2+4y^2+9z^2+1},\quad \frac{18z}{x^2+4y^2+9z^2+1}\right]^T.$$

Let us move on to the calculation of the second derivatives. Let's start with the calculation of the derivatives of the function $f'_x(x,y,z)$. We have to apply the quotient rule. The derivatives with respect to x, y and z are (in this order):

$$f''_{xx}(x,y,z)=\frac{2(x^2+4y^2+9z^2+1)-2x\cdot 2x}{\left(x^2+4y^2+9z^2+1\right)^2}=\frac{2(-x^2+4y^2+9z^2+1)}{\left(x^2+4y^2+9z^2+1\right)^2},$$

$$f''_{yx}(x,y,z)=\frac{0(x^2+4y^2+9z^2+1)-2x\cdot 8y}{\left(x^2+4y^2+9z^2+1\right)^2}=\frac{-16xy}{\left(x^2+4y^2+9z^2+1\right)^2},$$

$$f''_{zx}(x,y,z)=\frac{0(x^2+4y^2+9z^2+1)-2x\cdot 18z}{\left(x^2+4y^2+9z^2+1\right)^2}=\frac{-36xz}{\left(x^2+4y^2+9z^2+1\right)^2}.$$

In the same way we determine the derivatives of the functions $f'_y(x,y,z)$ and $f'_z(x,y,z)$. The Hessian is:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$Hf(x,y,z)=\begin{bmatrix} \dfrac{2(-x^2+4y^2+9z^2+1)}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{-16xy}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{-36xz}{\left(x^2+4y^2+9z^2+1\right)^2} \\ \dfrac{-16xy}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{8(x^2-4y^2+9z^2+1)}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{-144yz}{\left(x^2+4y^2+9z^2+1\right)^2} \\ \dfrac{-36xz}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{-144yz}{\left(x^2+4y^2+9z^2+1\right)^2} & \dfrac{18(x^2+4y^2-9z^2+1)}{\left(x^2+4y^2+9z^2+1\right)^2} \end{bmatrix}.$$

Note that also this time the respective mixed derivatives are equal, just as the Schwarz' Theorem says.

EXERCISES

1. Find the gradient and Hessian of the function:
   a) $f(x,y)=2x^2+3xy+5y^2-14x-26y$;
   b) $f(x,y)=3x^2-5xy+4y^2+2x+6y$;
   c) $f(x,y)=x^2e^{x+y}$;
   d) $f(x,y)=y^2e^{x-y}$;
   e) $f(x,y)=(x+y)e^{x^2}$;
   f) $f(x,y)=(x-y)e^{y^2}$;
   g) $f(x,y,z)=\ln(x^2+1)+\ln(z^2+2)+\ln(y^2+3)$;
   h) $f(x,y,z)=\ln(xyz)$.

SOLUTIONS

1. a) $\nabla f(x,y)=\begin{bmatrix}4x+3y-14\\3x+10y-26\end{bmatrix}$, $Hf(x,y)=\begin{bmatrix}4 & 3\\3 & 10\end{bmatrix}$;

   b) $\nabla f(x,y)=\begin{bmatrix}6x-5y+8\\-5x+8y+6\end{bmatrix}$; $Hf(x,y)=\begin{bmatrix}6 & -5\\-5 & 8\end{bmatrix}$;

   c) $\nabla f(x,y)=\begin{bmatrix}(x^2+2x)e^{x+y}\\x^2e^{x+y}\end{bmatrix}$,

   $Hf(x,y)=\begin{bmatrix}(x^2+4x+2)e^{x+y} & (x^2+2x)e^{x+y}\\(x^2+2x)e^{x+y} & x^2e^{x+y}\end{bmatrix}$;

   d) $\nabla f(x,y)=\begin{bmatrix}y^2e^{x-y}\\(-y^2+2y)e^{x-y}\end{bmatrix}$,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$Hf(x,y)=\begin{bmatrix} y^2e^{x-y} & (-y^2+2y)e^{x-y} \\ (-y^2+2y)e^{x-y} & (y^2-4y+2)e^{x-y} \end{bmatrix};$$

e) $$\nabla f(x,y)=\begin{bmatrix} (2x^2+2xy+1)e^{x^2} \\ e^{x^2} \end{bmatrix},$$

$$Hf(x,y)=\begin{bmatrix} 2(2x^3+2x^2y+3x+y)e^{x^2} & 2xe^{x^2} \\ 2xe^{x^2} & 0 \end{bmatrix};$$

f) $$\nabla f(x,y)=\begin{bmatrix} e^{y^2} \\ (2xy-2y^2-1)e^{y^2} \end{bmatrix},$$

$$Hf(x,y)=\begin{bmatrix} 0 & 2ye^{y^2} \\ 2ye^{y^2} & 2(2xy^2-2y^3+x-3y)e^{y^2} \end{bmatrix};$$

g) $$\nabla f(x,y,z)=\begin{bmatrix} \dfrac{2x}{x^2+1} \\ \dfrac{2y}{y^2+2} \\ \dfrac{2z}{z^2+3} \end{bmatrix},$$

$$Hf(x,y,z)=\begin{bmatrix} \dfrac{2(-x^2+1)}{\left(x^2+1\right)^2} & 0 & 0 \\ 0 & \dfrac{4}{\left(y^2+2\right)^2} & 0 \\ 0 & 0 & \dfrac{2(z^2+1)}{\left(z^2+3\right)^2} \end{bmatrix};$$

h) $$\nabla f(x,y,z)=\begin{bmatrix} \dfrac{1}{x} \\ \dfrac{1}{y} \\ \dfrac{1}{z} \end{bmatrix},\ Hf(x,y,z)=\begin{bmatrix} -\dfrac{1}{x^2} & 0 & 0 \\ 0 & -\dfrac{1}{y^2} & 0 \\ 0 & 0 & -\dfrac{1}{z^2} \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

9.2. LOCAL EXTREMA

THEORY IN A NUTSHELL

As in the case of functions of one variable, function of many variables can have local extrema. Unfortunately, it will be rather hard to imagine their graphical representation (you barely could draw a graph of the function of two variables – in three dimensions, for a function of three variables it is impossible). Remember, however, that intuitively we define them in the same way: it is the point at which the function value is greater (not less) or less (not more) than at the points in the closest neighborhood.
In the case of functions of many variables we can also distinguish two types of stationary points. However, because of the complexity, we will limit ourselves only to the type 1 stationary points. In the case of functions of one variable these are the points for which the first derivative is equal to 0, i.e. $f'(x) = 0$. In the case of function of may variables the respective condition has the following form:

Theorem 9.2. Necessary condition for a local extremum
If function f has a local extremum at point $(x_1, x_2, \ldots, x_n)$ and all its partial derivatives at this point are defined, then:

$$\nabla f(x_1, x_2, ..., x_n) = [0, 0, ..., 0]^T .$$

In Chapter 8 it has been said that in order to verify the existence of an extreme, you can check the sign of the second derivative at the stationary point. We will use a similar method in the case of functions of many variables. To this end, we have to introduce the notion of **leading principal submatrix**.

Definicja 9.5. Leading principal submatrix
The k-th leading principal submatrix of matrix A, denoted by $A^{(k)}$, is a matrix consisting of the elements of k first rows and columns of A.

The determinant of $A^{(k)}$, $\det(A^{(k)})$, is called the k-th **leading principal minor**.The sufficient condition for a local extremum is defined as follows:

Theorem 9.3. Sufficient condition for a local extremum
Let $x = (x_1, x_2, \ldots, x_n)$ be a type 1 stationary point of function f, at which all the second order derivatives are defined. Let us denote the Hessian of f at x by H. Then:

1. If $\det(H^{(i)}) > 0$ for $i = 1, 2, \ldots, n$, then f has local minimum at x.
2. If $(-1)^i \det(H^{(i)}) > 0$ for $i = 1, 2, \ldots, n$, then f has local maximum at x.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

So what you are supposed to do, if you want to confirm the existence of an extremum at some point? You have to find the values of all the leading principal minors of the Hessian matrix. If they are all positive, then there is a minimum at this point. If their signs alternate between negative and positive, then there is a maximum. **Note**: each of the above inequalilties can be replaced with the tight one, if the respective minor is fixed as 0 (i.e., its value does not depend on the values of the variables).
In summary, the procedure of determining local extrema is as follows:

1. Determine all the partial derivatives of the function and equate them to 0. The solutions of the resulting system of equations are type 1 stationary points. If the system of equations has no solutions, the function does not have the extrema[23].
2. Find all the second order derivatives and the Hessian matrix.
3. Substitute each stationary point to the Hessian and verify whether the sequence of leading principal minors satisfies the conditions given in Theorem 9.3. If so, then there is an extremum, otherwise not.

In order to find the local extrema of a function using **WolframAlpha®**, you can use the command "extrema", just as in the case of function of one variable. For example, in order to solve Example 3 below, you can type:

"extrema (x^2-6x+5)(y^2-2y-8)-4z^2".

EXAMPLES

Example 1
Find the local extrema of the function $f(x_1, x_2) = x_1^2 + 3x_1x_2 + 11x_2^2 - 5x_1 - 25x_2$.

Solution
Recall that the gradient of this function is (check the examples from the previous section):

$$\nabla f(x_1, x_2) = [2x_1 + 3x_2 - 5, \quad 3x_1 + 22x_2 - 25]^T.$$

In order to find the stationary points, we have to equate the derivatives to 0 and solve the resulting system of equations:

[23] Precisely speaking the function can have extrema also at the type 2 stationary points. Analyzing this type of functions, however, goes beyond the scope of this book.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\begin{cases} 2x_1 + 3x_2 - 5 = 0, \\ 3x_1 + 22x_2 - 25 = 0. \end{cases}$$

It is a system of two linear equations. Its solution is $x_1 = 1$, $x_2 = 1$. Hence f has one stationary point: (1, 1).
In order to verify the existence of an extremum at this point, we have to find the Hessian. From the previous section we know that it is equal to

$$Hf(x_1, x_2) = \begin{bmatrix} 2 & 3 \\ 3 & 22 \end{bmatrix},$$

so also

$$Hf(1, 1) = \begin{bmatrix} 2 & 3 \\ 3 & 22 \end{bmatrix}.$$

Since

$$\det\left(H^{(1)} f(1, 1)\right) = \det([2]) = 2 > 0$$

and

$$\det\left(H^{(2)} f(1, 1)\right) = \det\left(\begin{bmatrix} 2 & 3 \\ 3 & 22 \end{bmatrix}\right) = 35 > 0,$$

the function f has a local minimum at (1, 1). The value of f at this point is $f_{\min}(1, 1) = 1^2 + 3 \cdot 1 \cdot 1 + 11 \cdot 1^2 - 5 \cdot 1 - 25 \cdot 1 = -15$.

Example 2
Find the local extrema of the function

$$f(x, y) = e^y (8x + x^2 - 4y).$$

Solution
In the previous section we have found the gradient of f:

$$\nabla f(x, y) = [e^y (8 + 2x), \quad e^y (8x + x^2 - 4y - 4)]^T.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to find the local extrema we have to solve the following system of equations:

$$\begin{cases} e^y(8+2x)=0, \\ e^y(8x+x^2-4y-4)=0. \end{cases}$$

Let us begin with the first equation. Observe that since $e^y(8 + 2x) = 0$, at least one of the following conditions has to be satisfied: $e^y = 0$ or $8 + 2x = 0$. We know however that the exponential function takes only positive values, so the first expression cannot be 0. It follows that $8 + 2x = 0$, i.e. $x = –4$. We substitute this to the second equation, bearing in mind that the an exponential function takes only positive values. We obtain: $8\cdot(–4) + (–4)^2 – 4y – 4 = 0$, hence $y = –5$. As you can see, f has one stationary point $(–4, –5)$.
In order to verify the existence of an extremum, we have to check the Hessian. Recall that it is equal to:

$$Hf(x,y)=\begin{bmatrix} 2e^y & e^y(8+2x) \\ e^y(8+2x) & e^y(8x+x^2-4y-8) \end{bmatrix}.$$

By substituting the stationary point, we obtain:

$$Hf(-4,-5)=\begin{bmatrix} 2e^{-5} & 0 \\ 0 & -4e^{-5} \end{bmatrix}.$$

As we can see, f does not have an extremum at $(–4, –5)$. It follows from the fact that

$$\det([2e^{-5}]) = 2e^{-5} > 0$$

and

$$\det\left(\begin{bmatrix} 2e^{-5} & 0 \\ 0 & -4e^{-5} \end{bmatrix}\right) = -8e^{-10} < 0.$$

Finally, f has no local extrema.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
Find the local extrema of the function

$$f(x, y, z) = (x^2 - 6x + 5)(y^2 - 2y - 8) - 4z^2 .$$

Solution
The gradient of f has the form:

$$\nabla f(x, y, z) = [(2x - 6)(y^2 - 2y - 8), \quad (x^2 - 6x + 5)(2y - 2), \quad -8z]^T .$$

In order to find the stationary points, we have to solve the following system of equations:

$$\begin{cases} (2x - 6)(y^2 - 2y - 8) = 0, \\ (x^2 - 6x + 5)(2y - 2) = 0, \\ -8z = 0. \end{cases}$$

From the last one it follows that $z = 0$. The two remaining ones can be written in the form:

$$\begin{cases} (2x - 6)(y - 4)(y + 2) = 0, \\ (x - 5)(x - 1)(2y - 2) = 0. \end{cases} .$$

The first equation may be satisfied if and only if at least one of the expressions in the brackets is equal to 0. We consider three cases.
Case 1: $2x - 6 = 0$, i.e. $x = 3$. We substitute $x = 3$ to the second equation and we obtain $(3 - 5)(3 - 1)(2y - 2) = 0$, hence $y = 1$. This way we obtain first stationary point: (3, 1, 0) (remember that $z = 0$).
Case 2: $y - 4 = 0$, i.e. $y = 4$. We substitute $y = 4$ to the second equation and we obtain $(x - 5)(x - 1)(2 \cdot 4 - 2) = 0$, hence $x = 5$ or $x = 1$. This way we obtain two stationary points: (5, 4, 0) and (1, 4, 0).
Case 3: $y + 2 = 0$, i.e. $y = -2$. We substitute $y = -2$ to the second equation and we obtain $(x - 5)(x - 1)(2 \cdot (-2) - 2) = 0$, hence $x = 5$ or $x = 1$. This way we obtain two stationary points: (5, –2, 0) and (1, –2, 0).
Finally, f has five stationary points: (3, 1, 0), (5, 4, 0), (1, 4, 0), (5, –2, 0) and (1, –2, 0). We have to verify each of them.
Hessian has the general form:

$$Hf(x, y, z) = \begin{bmatrix} 2(y^2 - 2y - 8) & (2x - 6)(2y - 2) & 0 \\ (2x - 6)(2y - 2) & 2(x^2 - 6x + 5) & 0 \\ 0 & 0 & -8 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to verify the stationary points, we substitute each of them to the Hessian and check the signs of respective determinants:

$$Hf(3,1,0)=\begin{bmatrix}-18 & 0 & 0\\ 0 & -8 & 0\\ 0 & 0 & -8\end{bmatrix},$$

$$\det([-18])=-18<0,$$

$$\det\left(\begin{bmatrix}-18 & 0\\ 0 & -8\end{bmatrix}\right)=144>0,$$

$$\det\left(\begin{bmatrix}-16 & 0 & 0\\ 0 & -8 & 0\\ 0 & 0 & 0\end{bmatrix}\right)=0,$$

so at this point f has a local maximum, $f_{\max}(3, 1, 0) = 36$.

$$Hf(5,4,0)=\begin{bmatrix}0 & 24 & 0\\ 24 & 0 & 0\\ 0 & 0 & -8\end{bmatrix},$$

$$\det([0])=0,$$

$$\det\left(\begin{bmatrix}0 & 24\\ 24 & 0\end{bmatrix}\right)=-576<0,$$

$$\det\left(\begin{bmatrix}0 & 24 & 0\\ 24 & 0 & 0\\ 0 & 0 & -8\end{bmatrix}\right)=4608>0,$$

so there is no extremum at this point.

$$Hf(1,4,0)=\begin{bmatrix}0 & -24 & 0\\ -24 & 0 & 0\\ 0 & 0 & -8\end{bmatrix},$$

$$\det([0])=0,$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\det\left(\begin{bmatrix} 0 & -24 \\ -24 & 0 \end{bmatrix}\right) = -576 < 0,$$

$$\det\left(\begin{bmatrix} 0 & -24 & 0 \\ -24 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix}\right) = 4608 > 0,$$

so there is no extremum at this point.

$$Hf(5,-2,0) = \begin{bmatrix} 0 & -24 & 0 \\ -24 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix},$$

$$\det([0]) = 0,$$

$$\det\left(\begin{bmatrix} 0 & -24 \\ -24 & 0 \end{bmatrix}\right) = -576 < 0,$$

$$\det\left(\begin{bmatrix} 0 & -24 & 0 \\ -24 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix}\right) = 4608 > 0,$$

so there is no extremum at this point.

$$Hf(1,-2,0) = \begin{bmatrix} 0 & 24 & 0 \\ 24 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix},$$

$$\det([0]) = 0,$$

$$\det\left(\begin{bmatrix} 0 & 24 \\ 24 & 0 \end{bmatrix}\right) = -576 < 0,$$

$$\det\left(\begin{bmatrix} 0 & 24 & 0 \\ 24 & 0 & 0 \\ 0 & 0 & -8 \end{bmatrix}\right) = 4608 > 0,$$

so there is no extremum at this point.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXERCISES

1. Find the local extrema of the functions:
   a) $f(x,y) = 2x^2 + 3xy + 5y^2 - 14x - 26y$;
   b) $f(x,y) = 3x^2 - 5xy + 4y^2 + 2x + 6y$;
   c) $f(x,y) = 2x^2 - 3xy + 5y^2 + 2x + 14y$;
   d) $f(x,y) = x^2 - 4xy + 5y^2 + 4x - 12y$;
   e) $f(x,y) = (x+y)e^{x-y}$;
   f) $f(x,y) = (x+y)e^{y-x}$;
   g) $f(x,y,z) = \ln(3 - x^2 - y^2 - z^2)$;
   h) $f(x,y,z) = e^{1-x^2-y^2-z^2}$;
   i) $f(x,y,z) = 2\ln x + 3\ln y + 7\ln z + \ln(13 - x - y - z)$;
   j) $f(x,y,z) = \dfrac{1}{x^2 + 5y^2 + 3z^2 - 4xy + 2xz - 6yz + 2y - 8z}$.

SOLUTIONS

1. a) $f_{\min}(2, 2) = -40$; b) $f_{\min}(-2, -2) = -8$;
   c) $f_{\min}(-2, -2) = -16$; d) $f_{\min}(2, 2) = -8$;
   e) no stationary points, no extrema;
   f) no stationary points, no extrema;
   g) $f_{\max}(0, 0, 0) = \ln 3$; h) $f_{\max}(0, 0, 0) = e$;
   i) $f_{\max}(2, 3, 7) = 2 \ln 2 + 3 \ln 3 + 7 \ln 7$;
   j) $f_{\max}(1, 2, 3) = -1/10$.

9.3. CONDITIONAL EXTREMA

THEORY IN A NUTSHELL

Conditional extremum of a function is an extremum on a set defined by some set of constraints (equations and inequalities). This is the same kind of problem, as linear programming problem. The difference is that now we are interested in the nonlinear functions, so also the solution methods must change. Because of the complication level we will restrict ourselves to the problems with one constraints.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Assume that we want to find a minimum or maximum of the function $f(x_1, x_2, ..., x_n)$ subject to the constraint[24] $g(x_1, x_2, ..., x_n) = 0$. In order to find it, we construct an auxiliary function called the **Lagrange function**:

$$L(u, x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n) + ug(x_1, x_2, ..., x_n).$$

The new variable *u* is called **dual variable** or **Lagrange multiplier**[25].
Next we proceed as in the case of local extrema. To determine the stationary points, we equate to 0 the gradient of the Lagrange function.

Theorem 9.4. Necessary condition for a conditional extremum
If function $f(x_1, x_2, ..., x_n)$ has an extremum subject to constraint $g(x_1, x_2, ..., x_n) = 0$ at point $(x_1, x_2, \ldots, x_n)$ and all the partial derivatives of the Lagrange function are defined at this point, then

$$\nabla L(u, x_1, x_2, ..., x_n) = [0, 0, ..., 0]^T.$$

Note – **it is really important** – that the variable *u* is the first argument of *L* (i.e., it is placed **before** all the arguments of *f*). It does not matter when finding the stationary points, but it is **extremely important** during the verification using the Hessian. This verification is a bit different than in the case of local extrema, although very similar. It is described by the next theorem (remember that the dual variable *u* must be at the beginning!).

Theorem 9.5. Sufficient condition for a conditional extremum
Let $x = (u, x_1, x_2, \ldots, x_n)$ be a type 1 stationary point of function *L*, at which all the second order derivatives of *L* are defined. Let us denote the Hessian of *L* at *x* by *H*. Then:

1. If $\det(H^{(i)}) < 0$ for $i = 3, 4, \ldots, n + 1$, then *f* has conditional minimum at *x*.
2. If $(-1)^i \det(H^{(i)}) < 0$ for $i = 3, 4, \ldots, n + 1$, then *f* has conditional maximum at *x*.

Given conditions are different that in the case of local optima. Firstly, the numbering is shifted by 1, because there is an additional variable *u*. So the

[24] Note that the constraint has the form of equation. You should remember, however, that we can transform every inequality to the form of equation by introducing slack variables (see Chapter 5).
[25] In the case of larger number of constraints, we introduce one dual variable for each of them. As it was already told, however, in this book we restrict ourselves to the problems with one constraint.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Hessian has $n+1$ (rather than n) of rows and columns, and column corresponding with any variable x_j is the column $j+1$. Secondly, we are not interested in the signs of two first minors (we start only from the third leading principal submatrix). Thirdly, inequalities are reversed, and so we are interested in cases where the sequence of determinants is wholly negative or the elements are alternately positive and negative (the first one necessarily positive this time). Note below Theorem 9.2. also applies in this case.
If you want to find the conditional extrema of a function subject to a given constraint using **WolframAlpha®**, you can use the command "extrema" together with the expression "over" to indicate the constraint. For example, if you want to solve Example 2 below, you could type:

"extrema x1^2+x2^2 over 6x1+8x2=50".

EXAMPLES

Example 1
Find the extrema of the function

$$f(x,y)=(x-5)(y-6)$$

subject to

$$(x+7)(y-8)=0\,.$$

Solution
We begin with constructing the Lagrange function:

$$L(u,x,y)=(x-5)(y-6)+u(x+7)(y-8).$$

Its gradient is (remember that the first derivative is with respect to u, then we insert the derivatives with respect to x and y):

$$\nabla L(u,x,y)=\left[(x+7)(y-8),\quad (y-6)+u(y-8),\quad (x-5)+u(x+7)\right]^T.$$

It is necessary to solve the system of equations:

$$\begin{cases}(x+7)(y-8)=0,\\(y-6)+u(y-8)=0,\\(x-5)+u(x+7)=0.\end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

From the first equation it follows that $x = -7$ or $y = 8$. In the first case, by substituting $x = -7$ to the third equation we obtain: $(-7 - 5) + u(-7 + 7) = 0$, i.e. $-12 = 0$ – a contradiction. In the first case, by substituting $y = 8$ to the second equation we obtain $(8 - 6) + u(8 - 8) = 0$, i.e. $2 = 0$ – a contradiction, too. The system of equations has no solution, so the function has no stationary points and hence f has no conditional extrema.

Example 2
Find the extrema of the function

$$f(x_1, x_2) = x_1^2 + x_2^2$$

subject to

$$6x_1 + 8x_2 = 50.$$

Solution
The Lagrange function has the form (note that first, the condition must be transformed into a form having the right hand side equal to 0):

$$L(u, x_1, x_2) = x_1^2 + x_2^2 + u(6x_1 + 8x_2 - 50).$$

The gradient is:

$$\nabla L(u, x_1, x_2) = [6x_1 + 8x_2 - 50, \quad 2x_1 + 6u, \quad 2x_2 + 8u]^T.$$

It means that the following system must be solved:

$$\begin{cases} 6x_1 + 8x_2 - 50 = 0, \\ 2x_1 + 6u = 0, \\ 2x_2 + 8u = 0. \end{cases}$$

It is a linear system with solution: $u = -1$, $x_1 = 3$, $x_2 = 4$. Hessian of the Lagrange function equals to:

$$HL(-1, 3, 4) = \begin{bmatrix} 0 & 6 & 8 \\ 6 & 2 & 0 \\ 8 & 0 & 2 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Remember that in the case of conditional extrema we are checking only the signs of the leading principal minors starting from the third one. It means that in this example we are interested in only one determinant:

$$\det\left(\begin{bmatrix} 0 & 6 & 8 \\ 6 & 2 & 0 \\ 8 & 0 & 2 \end{bmatrix}\right) = -200 < 0 .$$

It means that the function f has a conditional minimum at (3, 4) (remember that the signs of the minors should be opposite to the signs in the case of the local minimum). The value of f at this point is $f_{\min}(3, 4) = 3^2 + 4^2 = 25$.

Example 3
Find the extrema of the function

$$f(x, y, z) = x^2 + y^2 + z^2$$

subject to

$$xyz = 125.$$

Solution
The Lagrange function has the form:

$$L(u, x, y, z) = x^2 + y^2 + z^2 + u(xyz - 125) .$$

Its gradient is:

$$\nabla L(u, x, y, z) = [xyz - 125, \quad 2x + uyz, \quad 2y + uxz, \quad 2z + uxy]^T .$$

It means that we have to solve the system:

$$\begin{cases} xyz - 125 = 0, \\ 2x + uyz = 0, \\ 2y + uxz = 0, \\ 2z + uxy = 0. \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Let us write it in a more convenient form:

$$\begin{cases} xyz = 125, \\ 2x = -uyz, \\ 2y = -uxz, \\ 2z = -uxy. \end{cases}$$

The system is such that it is convenient to divide the equation side-wise. To do this, however, you need first to make sure that the expressions on both sides do not divide by 0. Note that x, y, or z may not be equal to 0, because it would imply a contradiction in the first equation. It follows that also u cannot be zero (because then the remaining equations would imply that x, y and z are also zeros, and we have just shown that this is impossible). Divide the second equation by the third one. This way we obtain:

$$\frac{2x}{2y} = \frac{-uyz}{-uxz},$$

hence $y = x$ or $y = -x$. Doing the same with second and fourth equations, we conclude that $z = x$ or $z = -x$. Thus we have to consider four cases.
Case 1: $y = x$ and $z = x$. By substituting it to the first equation, we obtain $x^3 = 125$, i.e. $x = 5$, hence $y = 5$, $z = 5$ and (after substituting to any of the three last equations) $u = -2/5$.
Case 2: $y = x$ and $z = -x$. By substituting it to the first equation, we obtain $x^3 = -125$, i.e. $x = -5$, hence $y = -5$, $z = 5$ and (after substituting to any of the three last equations) $u = -2/5$.
Case 3: $y = -x$ and $z = x$. By substituting it to the first equation, we obtain $x^3 = -125$, i.e. $x = -5$, hence $y = 5$, $z = -5$ and (after substituting to any of the three last equations) $u = -2/5$.
Case 4: $y = -x$ and $z = -x$. By substituting it to the first equation, we obtain $x^3 = 125$, i.e. $x = 5$, hence $y = -5$, $z = -5$ and (after substituting to any of the three last equations) $u = -2/5$.
The Hessian has the form:

$$HL(u, x, y, z) = \begin{bmatrix} 0 & yz & xz & xy \\ yz & 2 & uz & uy \\ xz & uz & 2 & ux \\ xy & uy & ux & 2 \end{bmatrix}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The verification of the stationary points allows us to conclude that f has four conditional extrema (remember that we are interested only in the third and fourth leading principal minors here).

$$HL(-2/5,5,5,5)=\begin{bmatrix} 0 & 25 & 25 & 25 \\ 25 & 2 & -2 & -2 \\ 25 & -2 & 2 & -2 \\ 25 & -2 & -2 & 2 \end{bmatrix},$$

$$\det\left(\begin{bmatrix} 0 & 25 & 25 \\ 25 & 2 & -2 \\ 25 & -2 & 2 \end{bmatrix}\right)=-5\,000<0,$$

$$\det\left(\begin{bmatrix} 0 & 25 & 25 & 25 \\ 25 & 2 & -2 & -2 \\ 25 & -2 & 2 & -2 \\ 25 & -2 & -2 & 2 \end{bmatrix}\right)=-30\,000<0.$$

Hence $f_{\min}(5, 5, 5) = 75$.

$$HL(-2/5,-5,-5,5)=\begin{bmatrix} 0 & -25 & -25 & 25 \\ -25 & 2 & -2 & 2 \\ -25 & -2 & 2 & 2 \\ 25 & 2 & 2 & 2 \end{bmatrix},$$

$$\det\left(\begin{bmatrix} 0 & -25 & -25 \\ -25 & 2 & -2 \\ -25 & -2 & 2 \end{bmatrix}\right)=-5\,000<0,$$

$$\det\left(\begin{bmatrix} 0 & -25 & -25 & 25 \\ -25 & 2 & -2 & 2 \\ -25 & -2 & 2 & 2 \\ 25 & 2 & 2 & 2 \end{bmatrix}\right)=-30\,000<0.$$

Hence $f_{\min}(-5, -5, 5) = 75$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$HL(-2/5,-5,5,-5)=\begin{bmatrix} 0 & -25 & 25 & -25 \\ -25 & 2 & 2 & -2 \\ 25 & 2 & 2 & 2 \\ -25 & -2 & 2 & 2 \end{bmatrix},$$

$$\det\left(\begin{bmatrix} 0 & -25 & 25 \\ -25 & 2 & 2 \\ 25 & 2 & 2 \end{bmatrix}\right)=-5\,000<0,$$

$$\det\left(\begin{bmatrix} 0 & -25 & 25 & -25 \\ -25 & 2 & 2 & -2 \\ 25 & 2 & 2 & 2 \\ -25 & -2 & 2 & 2 \end{bmatrix}\right)=-30\,000<0.$$

Hence $f_{\min}(-5, 5, -5) = 75$.

$$HL(-2/5,5,-5,-5)=\begin{bmatrix} 0 & 25 & -25 & -25 \\ 25 & 2 & 2 & 2 \\ -25 & 2 & 2 & -2 \\ -25 & 2 & -2 & 2 \end{bmatrix},$$

$$\det\left(\begin{bmatrix} 0 & 25 & -25 \\ 25 & 2 & 2 \\ -25 & 2 & 2 \end{bmatrix}\right)=-5\,000<0,$$

$$\det\left(\begin{bmatrix} 0 & 25 & -25 & -25 \\ 25 & 2 & 2 & 2 \\ -25 & 2 & 2 & -2 \\ -25 & 2 & -2 & 2 \end{bmatrix}\right)=-30\,000<0.$$

Hence $f_{\min}(5, -5, -5) = 75$.

EXERCISES

1. Find the extrema of the functions:
   a) $f(x,y)=(x-11)(y+13)$, subject to $(x+7)(y-8)=0$;
   b) $f(x,y)=(x+12)(y-3)$, subject to $(x+7)(y-8)=0$;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) $f(x, y) = x^2 + y^2$, subject to $x + y = 12$;

d) $f(x, y) = x^2 + y^2 + z^2$, subject to $x + y + z = 24$;

e) $f(x, y) = xy$, subject to $\ln(x + y) = 0$;

f) $f(x, y) = xy$, subject to $\ln(x^2 + y^2) = 0$;

g) $f(x, y, z) = xyz$, subject to $\ln(x + y + z) = 0$;

h) $f(x, y, z) = x + y + z$, subject to $\ln(x^2 + y^2 + z^2) = 0$;

i) $f(x, y) = e^{x^2+y^2}$, subject to $xy = 1$;

j) $f(x, y, z) = \ln(x^2 + y^2 + z^2)$, subject to $xyz = 1$.

SOLUTIONS

1. a) no stationary points, no extrema;

b) no stationary points, no extrema;

c) $f_{\min}(6, 6) = 72$;

d) $f_{\min}(8, 8, 8) = 192$;

e) $f_{\max}(1/2, 1/2) = 1/4$;

f) $f_{\min}\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\frac{1}{2}$, $f_{\min}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = -\frac{1}{2}$,
$f_{\max}\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \frac{1}{2}$, $f_{\max}\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \frac{1}{2}$;

g) $f_{\max}\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{27}$;

h) $f_{\max}\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right) = \sqrt{3}$, $f_{\min}\left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right) = -\sqrt{3}$;

i) $f_{\min}(1, 1) = e^2$, $f_{\min}(-1, -1) = e^2$;

j) $f_{\min}(1, 1, 1) = \ln 3$, $f_{\min}(1, -1, -1) = \ln 3$,
$f_{\min}(-1, 1, -1) = \ln 3$, $f_{\min}(-1, -1, 1) = \ln 3$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 10
INTEGRAL CALCULUS

10.1. INDEFINITE INTEGRALS

THEORY IN A NUTSHELL

Integration is the inverse of differentiation – to determine the integral of function *f*, you need to find a function whose derivative is *f*.

Definition 10.1. Antiderivative
$F(x)$ is antiderivative of $f(x)$, when $F'(x) = f(x)$.

Definition 10.2. Indefinite integral
Indefinite integral is the family of all antiderivatives of $f(x)$, differing by a constant: $\int f(x)dx = F(x) + C$.

Antiderivatives of chosen functions have been listed in Table 10.1. They correspond with the formulae for derivatives that you already know (see Section 8.2).

Table 10.1. Antiderivatives

$f(x)$	$F(x)$	$f(x)$	$F(x)$
0	0	$1/x$	$\ln\|x\|$
1	x	e^x	e^x
$x^n, n \neq -1$	$\frac{x^{n+1}}{n+1}$	a^x	$\frac{a^x}{\ln a}$

As in the case of derivatives, this list is not complete. In order to be able to find more integrals, you have to know their properties:

Theorem 10.1. Basic properties of indefinite integrals

1. Sum rule: $\int [c_1 f(x) + c_2 g(x)]dx = c_1 \int f(x)\,dx + c_2 \int g(x)\,dx.$
2. Substitution rule: $\int f(g(x))g'(x)dx = \int f(t)dt$, where $t = g(x)$, $dt = g'(x)dx.$
3. Integration by parts: $\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to find an integral using **WolframAlpha®**, you can apply the command "integrate". For example, if you want to solve Example 1 below, you can type:

"integrate 12x^4+x^2-3x".

EXAMPLES

Example 1
Find the integral:

$$\int (12x^4 + x^2 - 3x)\,dx\,.$$

Solution
We can use directly the formulae:

$$\int (12x^4 + x^2 - 3x)\,dx = 12\frac{x^{4+1}}{4+1} + \frac{x^{2+1}}{2+1} - 3\frac{x^{1+1}}{1+1} + C = \frac{12}{5}x^5 + \frac{1}{3}x^3 - \frac{3}{2}x^2 + C\,.$$

Example 2
Find the integral:

$$\int \frac{7x\sqrt{x} - 11x\sqrt[3]{x}}{\sqrt[6]{x}}\,dx\,.$$

Solution
We begin with some transformations – we express the roots in the form of powers and then use some properties of powers:

$$\int \frac{7x\sqrt{x} - 11x\sqrt[3]{x}}{\sqrt[6]{x}}\,dx = \int \frac{7x^{3/2} - 11x^{4/3}}{x^{1/6}}\,dx = \int (7x^{4/3} - 11x^{7/6})\,dx = \ldots$$

Now we can apply directly the formulae:

$$\ldots = 7\frac{x^{4/3+1}}{4/3+1} - 11\frac{x^{7/6+1}}{7/6+1} + C = 3x^{7/3} - \frac{66}{13}x^{13/6} + C\,.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
Find the integral:

$$\int 7^x 11^x \, dx .$$

Solution
We can find the antiderivative by using the formula for exponential function. First, we have to transform the function to the form of a power with exponent x:

$$\int 7^x 11^x \, dx = \int (7 \cdot 11)^x \, dx = \int 77^x \, dx = \frac{77^x}{\ln 77} + C .$$

Example 4
Find the integral:

$$\int e^{3x} dx .$$

Solution
Similarly as in the previous example, we can transform the function to the power with exponent x (using the fact that $e^{3x} = (e^3)^x$). However, we will use another method. We are going to apply the substitution rule. As there is $3x$ in the exponent (and there is x in the formula), we will substitute it with new variable t. Next step is finding the derivative of t with respect to x – recall that it can be written as dt / dx. Our last goal is to find dx, in order to be able to substitute it with some expression containing t.

$$\int e^{3x} dx = \left| \begin{array}{c} t = 3x \\ dt / dx = 3 \\ dx = 1/3 dt \end{array} \right| = \ldots$$

We can now substitute t for $3x$ and $1/3dt$ for dx. This way we obtain an integral that can be derived by using the standard formulae (do not worry that the variable has changed – it is just a letter):

$$\ldots = \int e^t \cdot 1/3 dt = 1/3 e^t + C = \ldots$$

Finally, we can go back to the previous notation:

$$\ldots = 1/3 e^{3x} + C .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 5
Find the integral:

$$\int \left(x^2+3\right)^{30} x\,dx\,.$$

Solution
In order to find this integral, you can start with multiplying the expression in brackets. However, repeating it thirty times would take much time and it would be very likely that you make a mistake. For that reason we will use the substitution rule – we will substitute *t* for the expression in brackets:

$$\int \left(x^2+3\right)^{30} x\,dx = \left| \begin{array}{c} t = x^2+3 \\ dt/dx = 2x \\ \ldots \end{array} \right| .$$

This time it is impossible to transform the equation in such way that there would be *dx* on the left and only expressions with *t* on the right. But it is not a problem – it just happens so that we need an expression with $x\,dx$. The remainder of the reasoning is elementary:

$$\int \left(x^2+3\right)^{30} x\,dx = \left| \begin{array}{c} t = x^2+3 \\ dt/dx = 2x \\ xdx = 1/2dt \end{array} \right| = \int t^{30} \cdot 1/2 dt = \frac{t^{31}}{62} + C = \frac{\left(x^2+3\right)^{31}}{62} + C\,.$$

Example 6
Find the integral:

$$\int x^4 e^{x^5}\,dx\,.$$

Solution
This time we have to substitute the exponent. Also this time we do not try to express *dx* in terms of *t* – we express together $x^4 dx$:

$$\int x^4 e^{x^5} dx = \left| \begin{array}{c} t = x^5 \\ dt/dx = 5x^4 \\ x^4 dx = 1/5 dt \end{array} \right| = \int e^t \cdot 1/5 dt = 1/5 e^t + C = 1/5 e^{x^5} + C\,.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 7
Find the integral:

$$\int \frac{\ln^7 x}{x} dx .$$

Solution
This time we substitute t for $\ln x$. Thanks to that we can use the formula for the power function.

$$\int \frac{\ln^7 x}{x} dx = \left| \begin{array}{r} t = \ln x \\ dt/dx = 1/x \\ dt = dx/x \end{array} \right| = \int t^7 dt = \frac{t^8}{8} + C = \frac{\ln^8 x}{8} + C .$$

Example 8
Find the integral:

$$\int x \cdot 13^x dx .$$

Solution
This time we deal with a product of two functions of different kinds (power and exponential). In such situation we shall apply the integration by parts. In this case we differentiate the power function (i.e., x) and find the antiderivative of the exponential function.

$$\int x \cdot 13^x dx = \begin{vmatrix} u(x) = x & v'(x) = 13^x \\ \ldots & \ldots \end{vmatrix} .$$

The derivative of x is 1, and the sample antiderivative of 13^x is $13^x / \ln 13$. We integrate by parts:

$$\int x \cdot 13^x dx = \begin{vmatrix} u(x) = x & v'(x) = 13^x \\ u'(x) = 1 & v(x) = 13^x / \ln 13 \end{vmatrix} = x \cdot \frac{13^x}{\ln 13} - \int 1 \cdot \frac{13^x}{\ln 13} dx = \ldots$$

As you can see, the last integral is simpler than the initial one – we can find it using the standard formulae. Thus finally we obtain:

$$\ldots = x \cdot \frac{13^x}{\ln 13} - \frac{13^x}{(\ln 13)^2} + C .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 9
Find the integral:

$$\int \ln x \, dx .$$

Solution
The integrals of this kind we also solve by parts. It may seem strange, as you see only one function. Remember however, that you can always multiply an expression by 1, not changing its value. If an expression with logarithm has to be integrated, then usually the logarithm will be differentiated, as in this example.

$$\int \ln x \, dx = \int 1 \cdot \ln x \, dx = \begin{vmatrix} u(x) = \ln x & v'(x) = 1 \\ u'(x) = 1/x & v(x) = x \end{vmatrix} =$$

$$= (\ln x) \cdot x - \int 1/x \cdot x \, dx = x \ln x - x + C.$$

Example 10
Find the integral:

$$\int x^8 e^{x^3} dx .$$

Solution
In this case we begin with a substitution (you always need to substitute for the expression in the exponent). Note however that there is also the expression x^8. It is necessary to transform it in order to factor out x^3 (i.e. the expression, for which we substitute).

$$\int x^8 e^{x^3} dx = \int \left(x^3\right)^2 x^2 e^{x^3} dx = \begin{vmatrix} t = x^3 \\ dt / dx = 3x^2 \\ x^2 dx = 1/3 dt \end{vmatrix} = \int t^2 e^t \cdot 1/3 dt = 1/3 \int t^2 e^t dt = \ldots$$

We obtained integral that should be solved by parts. Just as in Example 8, we differentiate the power function and find the antiderivative of the exponential part.

$$\ldots = \begin{vmatrix} u(t) = t^2 & v'(t) = e^t \\ u'(t) = 2t & v(t) = e^t \end{vmatrix} = 1/3 \left(t^2 e^t - \int 2te^t dt \right) = 1/3 t^2 e^t - 2/3 \int te^t dt = \ldots$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

We solve the last integral by parts and go back to the initial notation.

$$\ldots = \begin{vmatrix} u(t)=t & v'(t)=e^t \\ u'(t)=1 & v(t)=e^t \end{vmatrix} = 1/3t^2e^t - 2/3(te^t - \int e^t dt) =$$

$$= 1/3t^2e^t - 2/3te^t + 2/3e^t + C = 1/3x^6e^{x^3} - 2/3x^3e^{x^3} + 2/3e^{x^3} + C\,.$$

EXERCISES

1. Find the integrals, using the basic formulae and transformations:
   a) $\int \frac{5x^2\sqrt{x} + 7x^2\sqrt[3]{x}}{\sqrt[6]{x^5}}dx$;
   b) $\int \frac{17x^2\sqrt{x} - 11x^3\sqrt[3]{x}}{x\sqrt[3]{x^2}}dx$;
   c) $\int 5^x 13^x dx$;
   d) $\int 5^{3x} 2^{2x} dx$;
   e) $\int \frac{4dx}{x^3}$.
2. Find the integrals by applying the substitution rule:
   a) $\int \left(x^3+5\right)^{21} x^2 dx$;
   b) $\int \left(x^4-2\right)^{19} x^3 dx$;
   c) $\int \frac{dx}{(x-2)}$;
   d) $\int \frac{dx}{(x+3)^6}$;
   e) $\int e^{4x} dx$;
   f) $\int e^{-5x+17} dx$;
   g) $\int e^{x^3} x^2 dx$;
   h) $\int e^{5x^5} x^4 dx$;
   i) $\int e^{3x^2+2} x\,dx$;
   j) $\int \frac{\ln^{11} x}{x} dx$;
   k) $\int \frac{\ln(2x+3)}{2x+3} dx$;
   l) $\int \frac{\ln^3(3e^x+7)}{3e^x+7} e^x dx$.
3. Find the integrals by applying the integration by parts (in some cases it will be necessary to apply also the substitution rule):
   a) $\int xe^x dx$;
   b) $\int x^2 e^{7x} dx$;
   c) $\int 5xe^{-5x} dx$;
   d) $\int x^5 e^{x^3} dx$;
   e) $\int x^{11} e^{x^4} dx$;
   f) $\int \ln^3 x\, dx$;
   g) $\int x^{-7} \ln x\, dx$;
   h) $\int \ln(x^5) dx$;
   i) $\int x \ln(x^3) dx$;
   j) $\int \frac{\ln(x^5)}{x^6} dx$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. a) $15/8x^{8/3} + 14/5x^{5/2} + C$; b) $102/11x^{11/6} - 33/8x^{8/3} + C$;
   c) $65^x/\ln 65 + C$; d) $500^x/\ln 500 + C$;
   e) $-2/(x^2) + C$.
2. a) $(x^3 + 5)^{22}/66 + C$; b) $(x^4 - 2)^{20} / 80 + C$;
   c) $\ln|x - 2| + C$; d) $-1/[5(x + 3)^5] + C$;
   e) $e^{4x}/4 + C$; f) $e^{-5x + 17}/(-5) + C$;
   g) $e^{x^3}/3 + C$; h) $e^{5x^5}/25 + C$;
   i) $e^{3x^2+2}/6 + C$; j) $\ln^{12}x/12 + C$;
   k) $\ln^2(2x + 3)/4 + C$; l) $\ln^4(3e^x + 7)/12 + C$.
3. a) $xe^x - e^x + C$;
   b) $1/7x^2e^{7x} - 2/49xe^{7x} + 2/343e^{7x} + C$;
   c) $-xe^{-5x} - 1/5e^{-5x} + C$;
   d) $1/3x^3e^{x^3} - 1/3x^3 + C$;
   e) $1/4x^8e^{x^4} - 1/2x^4e^{x^4} + 1/2e^{x^4} + C$;
   f) $x\ln^3 x - 3x\ln^2 x + 6x\ln x - 6x + C$;
   g) $-1/6x^{-6}\ln x - 1/36x^{-6} + C$;
   h) $5x\ln x - 5x + C$;
   i) $3/2x^2\ln x - 3/4x^2 + C$;
   j) $-x^{-5}\ln x - 1/5x^{-5} + C$.

10.2. DEFINITE AND IMPROPER INTEGRALS

Definition 10.3. Definite integral

Definite integral of function $f(x)$ over the interval $[a, b]$ is the difference of the values of any chosen antiderivative of $f(x)$ at b and a:

$$\int_a^b f(x)dx = \left[F(x)\right]_a^b = F(b) - F(a).$$

In the case of definite integrals we can also provide a set of useful properties.

Theorem 10.2. Basic properties of definite integrals

1. Sum rule: $\int_a^b [c_1 f(x) + c_2 g(x)]dx = c_1\int_a^b f(x)dx + c_2\int_a^b g(x)dx$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Substitution rule: $\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt$, where $t = g(x)$, $dt = g'(x)dx$.
3. Integration by parts: $\int_a^b u(x)v'(x)dx = \left[u(x)v(x)\right]_a^b - \int_a^b u'(x)v(x)dx$.
4. Additivity of integration on intervals: $\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx$.

Definite integral has a simple graphical interpretation: it is equal to the area of the region between the graph of the integrand and the x-axis. For example, the area of the hatched region in Fig. 10.1a is equal to

$$S_{reg} = \int_a^b f(x)dx.$$

It can be generalized to the regions bounded by two functions. If a region is bounded on the interval $[a, b]$ by function $f(x)$ from above, and by $g(x)$ from below (just as in Fig. 10.1b), then the area of this region equals to:

$$S_{reg} = \int_a^b (f(x) - g(x))dx.$$

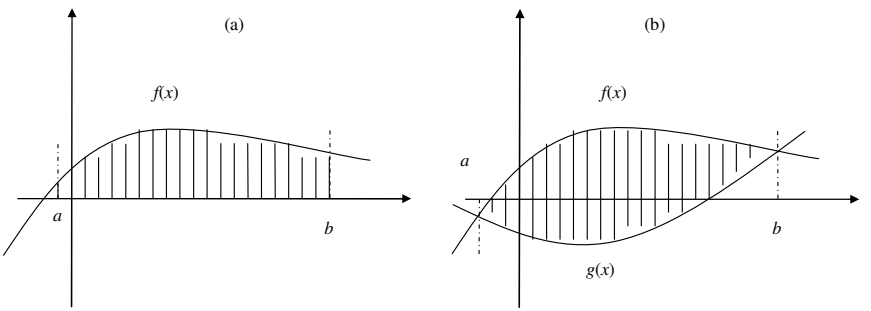


Fig. 10.1. Geometric interpretation of definite integrals

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Sometimes it happens that the integrand is not defined at some point of the interval of integration or that this interval is unbounded. Such a special integral is called **improper integral**.

Definition 10.4. Improper integral
Improper integral is the limit of a definite integral defined on an unbounded interval or the definite integral of a function undefined at some point of the interval of integration.

In order to **verify the convergence** of an improper integral, it is necessary to calculate the limit as x approaches the endpoint of the interval of integration equal to $-\infty$ or ∞, or x approaches the point at which the integrand is not defined. If it is necessary, the domain of integration should be divided into intervals in such way that each limit is calculated at the end of some interval and there is only one limit to find in each of those intervals. If all the resulting integrals are finite, then the initial integral is their sum. Otherwise the initial integral is **divergent**.
In order to find the value of a definite integral or test the convergence of an improper one using **WolframAlpha®**, you can apply the command "integrate", defining the domain of integration with the words "from" and "to". For example, if you want to solve Example 1 below, you can type:

"integrate x^2-7x+3 from 1 to 5",

while if you want to test the convergence of the integral from Example 7, you should type:

"integrate 1/x^2 from -inf to inf".

In order to find the area between curves, you can use the command "area between". For example, if you want to solve Example 5, just type:

"area between y=2x+1, y=x^2-x+3".

EXAMPLES

Example 1
Find the value of definite integral:

$$\int_1^5 (x^2 - 7x + 3)dx .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
We use directly the formula (the function in square brackets is an antiderivative of the integrand – note that we skip the constant *C*):

$$\int_1^5 (x^2 - 7x + 3)dx = \left[\frac{x^3}{3} - \frac{7x^2}{2} + 3x\right]_1^5 = \ldots$$

Now we compute the values of the antiderivative at both ends of the domain of integration and subtract:

$$\ldots = \left[\frac{5^3}{3} - \frac{7 \cdot 5^2}{2} + 3 \cdot 5\right] - \left[\frac{1^3}{3} - \frac{7 \cdot 1^2}{2} + 3 \cdot 1\right] = -92/3 .$$

Example 2
Find the value of definite integral:

$$\int_0^{20} |x - 12| dx .$$

Solution
The integrand has the form of absolute value of an expression which is negative for $x < 12$ and nonnegative for $x \geq 12$. Hence it can be written as:

$$|x - 12| = \begin{cases} -x + 12, \text{ if } x < 12, \\ x - 12, \quad \text{if } x \geq 12. \end{cases}$$

Thus we can use the additivity of integration on intervals and write the integral in the form:

$$\int_0^{20} |x - 12| dx = \int_0^{12} (-x + 12)dx + \int_{12}^{20} (x - 12)dx = \ldots$$

Finally, we apply the respective formulae:

$$\ldots = \left[-1/2x^2 + 12x\right]_0^{12} + \left[1/2x^2 - 12x\right]_{12}^{20} = 72 + 32 = 104 .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 3
Find the value of definite integral:

$$\int_{e}^{e^4} \frac{dx}{x \ln x}.$$

Solution
This time we use the substitution rule. We substitute t for $\ln x$, so the ends of the domain of integration will be $\ln e = \ln e^1 = 1$ and $\ln e^4 = 4$, respectively.

$$\int_{e}^{e^4} \frac{dx}{x \ln x} = \left| \begin{array}{c} t = \ln x \\ dt / dx = 1/x \\ dt = dx / x \end{array} \quad \begin{array}{c|c} x & t = \ln x \\ \hline e & 1 \\ e^4 & 4 \end{array} \right| = \int_{1}^{4} \frac{dt}{t} = \left[\ln|t|\right]_1^4 = \ln 4 - \ln 1 = \ln 4 - 0 = \ln 4.$$

Example 4
Determine the area of the region enclosed by:

$$\begin{cases} y = 2x + 1, \\ y = x^2 - x + 3. \end{cases}$$

Solution
Let us begin with sketching the graphs of both functions. We will use the notation: $f(x) = 2x + 1$, $g(x) = x^2 - x + 3$ (the scale is intentionally not preserved, it is supposed to be only a sketch).

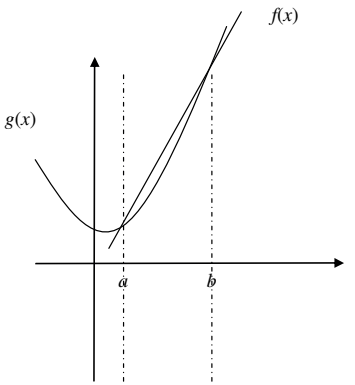


Fig. 10.2. Region enclosed by $f(x)$ and $g(x)$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to find the x-coordinates (i.e., a and b) of the intersection points of $f(x)$ and $g(x)$, we equate the functions: $2x + 1 = x^2 - x + 3$, hence $x^2 - 3x + 2 = 0$, i.e. $x = 1 \vee x = 2$. Thus $a = 1$ and $b = 2$. Thus the area of the enclosed region is equal to:

$$S_{reg} = \int_a^b (f(x) - g(x))dx = \int_1^2 [(2x+1) - (x^2 - x + 3)]dx = \left[-\frac{x^3}{3} + \frac{3}{2}x^2 - 2x \right]_1^2 = 1/6 .$$

Example 5
Examine the convergence of the integral:

$$\int_0^5 \frac{1}{4\sqrt{x}} dx .$$

Solution
It would be an ordinary definite integral, if not the fact that the integrand is not defined at the left end of the domain of integration, i.e. at $x = 0$. Because of that we have to determine the limit at 0 from above (0 is the left end, hence the integration domain is to the right of 0).

$$\int_0^5 \frac{1}{4\sqrt{x}} dx = \lim_{a \to 0^+} \int_a^5 \frac{1}{4\sqrt{x}} dx = ...$$

Now we find the integral with parameter a, bearing in mind that there is a limit to determine:

$$... = \lim_{a \to 0^+} \int_a^5 1/4 x^{-1/2} dx = \lim_{a \to 0^+} \left[1/2 x^{1/2} \right]_a^5 = \lim_{a \to 0^+} (1/2\sqrt{5} - 1/2\sqrt{a}) = \frac{\sqrt{5}}{2} .$$

Example 6
Examine the convergence of the integral:

$$\int_7^{11} \frac{dx}{x-9} .$$

Solution
In this case the integrand is not defined at $x = 9$, which is inside the domain of integration. For that reason we subdivide the domain, using the additivity of integration on intervals:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\int_7^{11}\frac{dx}{x-9}=\int_7^{9}\frac{dx}{x-9}+\int_9^{11}\frac{dx}{x-9}.$$

Note: we do not keep writing "=", instead of that we examine the convergence of each integral separately. Only if it turns out that they all converge, we add up their values.

$$\int_7^{9}\frac{dx}{x-9}=\lim_{a\to 9^-}\int_7^{a}\frac{dx}{x-9}=\lim_{a\to 9^-}\left[\ln|x-9|\right]_7^a=$$
$$=\lim_{a\to 9^-}\left(\ln|a-9|-\ln|7-9|\right)=[\ln 0^+-\ln 2]=-\infty.$$

Thus the integral $\int_7^{9}\frac{dx}{x-9}$ is divergent. It means that also $\int_7^{11}\frac{dx}{x-9}$ is divergent (the examination of convergence of $\int_9^{11}\frac{dx}{x-9}$ is not necessary).

Example 7
Examine the convergence of the integral:

$$\int_{-\infty}^{\infty}\frac{dx}{x^2}.$$

Solution
First, we have to take under consideration the fact that the integrand is not defined at $x = 0$. Thus we have to subdivide the domain of integration:

$$\int_{-\infty}^{\infty}\frac{dx}{x^2}=\int_{-\infty}^{0}\frac{dx}{x^2}+\int_0^{\infty}\frac{dx}{x^2}.$$

In case of both resulting integrals it is necessary to find two limits (at both ends of the domain of integration): in the first case – the limits at $-\infty$ and at 0^-, while in the second case – the limits at 0^+ and at ∞. It means that the domains of integration of both integrals must be further subdivided. As the limits will be computed at the ends of the domains of integration, and not inside it, the subdivision point can be arbitrary, provided that it is inside the domain. For example we can assume that the first interval will be subdivided at $x = -1$, and the second one at $x = 1$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\int_{-\infty}^{\infty}\frac{dx}{x^2}=\int_{-\infty}^{-1}\frac{dx}{x^2}+\int_{-1}^{0}\frac{dx}{x^2}+\int_{0}^{1}\frac{dx}{x^2}+\int_{1}^{\infty}\frac{dx}{x^2}.$$

Now we can proceed to the main part of the solution. Let us begin with examining the first integral.

$$\int_{-\infty}^{-1}\frac{dx}{x^2}=\lim_{a\to-\infty}\int_{a}^{-1}\frac{dx}{x^2}=\lim_{a\to-\infty}\int_{a}^{-1}x^{-2}dx=\lim_{a\to-\infty}\left[\frac{x^{-1}}{-1}\right]_{a}^{-1}=\lim_{a\to-\infty}\left[\frac{-1}{x}\right]_{a}^{-1}=\lim_{a\to-\infty}\left(\frac{-1}{-1}-\frac{-1}{a}\right)=1.$$

As you can see, it turned out to be convergent. Therefore, we proceed to test the convergence of the second one.

$$\int_{-1}^{0}\frac{dx}{x^2}=\lim_{a\to0^-}\int_{-1}^{a}\frac{dx}{x^2}=\lim_{a\to0^-}\left[\frac{-1}{x}\right]_{-1}^{a}=\lim_{a\to0^-}\left(\frac{-1}{a}-\frac{-1}{-1}\right)=\infty.$$

The second integral is divergent. It means that $\int_{-\infty}^{\infty}\frac{dx}{x^2}$ is also divergent.

EXERCISES

1. Find the values of definite integrals:
   a) $\int_{2}^{4}(x^3+5)dx$;
   b) $\int_{1}^{e^5}\ln x\,dx$;
   c) $\int_{-2}^{8}|2x-7|dx$;
   d) $\int_{0}^{4}|x^2-4x+3|dx$;
   e) $\int_{0}^{1}xe^{3x^2}dx$.
2. Determine the area of the region enclosed by:
   a) $y = x^2 - 6x + 10$, $y = 2x - 5$;
   b) $y = x^2 - 4x - 8$, $y = -2x + 7$;
   c) $y = x^2 - 10x + 15$, $y = -x^2 + 2x + 5$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

3. Examine the convergence of the integrals:

a) $\int_1^{\infty} \frac{4}{x^5} dx$; b) $\int_0^{36} \frac{2dx}{x^3}$;

c) $\int_0^{36} \frac{2dx}{\sqrt{x}}$; d) $\int_{-7}^{11} \frac{dx}{x}$;

e) $\int_{-\infty}^{\infty} \frac{dx}{x-11}$.

SOLUTIONS

1. a) 70; b) $4e^5 + 1$; c) 101/2; d) 4; e) $1/6e^3 - 1/6$.
2. a) 4/3; b) 256/3; c) 64/3.
3. a) 1; b) divergent; c) 24; d) divergent; e) divergent.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 11
DIFFERENTIAL AND DIFFERENCE EQUATIONS

11.1. DIFFERENTIAL EQUATIONS

THEORY IN A NUTSHELL

Definition 11.1. Ordinary differential equation
Ordinary differential equation is any equation relating a function of one variable $f(x)$ with at least one of its derivatives. The function $f(x)$ is unknown in the equation.

Definition 11.2. First order differential equation
First order differential equation is a differential equation relating $f(x)$ only with its first derivative and its argument. It can be written in the form:

$$R\left(\frac{dy}{dx}, y, x\right) = 0,$$

where $y = f(x)$.

Definition 11.3. Integral of differential equation
The solution of a differential equation, i.e. the function that satisfies it, is called its integral. **General integral** of a differential equation is the family of all the functions that satisfy it. **Particular integral** of a differential equation is the solution determined under assumption that the plot of the function contains some predefined point (x_0, y_0).

Particular integral is the solution of an **initial value problem**.

Definition 11.4. Initial value problem
Initial value problem (also called the **Cauchy problem**) is a system of two equations: ordinary differential equation and so called initial condition, i.e. the equation defining the value of the unknown function at a given point.

$$\begin{cases} R\left(\frac{dy}{dx}, y, x\right) = 0, \\ f(x_0) = y_0. \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In the next part we will focus on three types of differential equations: **equations with separated variables**, equations solved by **constant coefficient substitution** and **homogeneous equations**.

Definition 11.5. Equation with separated variables
Differential equation with separated variables is an equation that, using only algebraic transformations, can be reduced to a form in which on one side there are only expressions with *x*, and on the other expressions with *y*, i.e. equation that can be written as:

$$f(y)dy = g(x)dx.$$

Differential equations with separated variables are solved by integration of both sides.

Definition 11.6. Equation solved by constant coefficient substitution
Equation solved by constant coefficient substitution is an equation in which *x* and *y* are present only together, in the form of linear expression $ax+by+c$.

In order to solve an equation of this kind we use the substitution:

$$u = ax+by+c.$$

Hence

$$y = \frac{u-ax-c}{b},$$

$$\frac{dy}{dx} = \frac{1}{b}\cdot\frac{du}{dx} - \frac{a}{b}.$$

Definition 11.7. Homogeneous equation
Homogeneous equation is an equation in which *x* and *y* are present only together, in the form of the fraction $\frac{y}{x}$.

In order to solve an equation of this kind we use the substitution:

$$u = \frac{y}{x}.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Hence

$y = ux,$

$\frac{dy}{dx} = \frac{du}{dx}x + u\,.$

In order to solve a differential equation or an initial value problem using **WolframAlpha®**, just type it, optionally using a comma. For example, if you want to solve Example 2 below, type:

"7x^2 dy/dx = y",

and if you want to solve Example 5, type:

"dy/dx=12x, y(2)=33"

(note that we use the notation "$y(x)$" instead of "$f(x)$").

EXAMPLES

Example 1
Solve the differential equation:

$\frac{dy}{dx} = e^x\,.$

Solution
We transform the equation, so that the left side consists only of the expressions with y and the right side – only of the expressions with x:

$dy = e^x dx\,.$

We integrate both sides (we write the constant only on one side):

$\int dy = \int e^x dx\,,$

hence

$y = e^x + C\,.$

This way we obtained the final solution.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 2
Solve the differential equation:

$$7x^2 \frac{dy}{dx} = y .$$

Solution
We begin, as in the previous example, with moving the expressions with y to the left side, and the expressions with x to the right, and then we perform integration. This time, we divide by y, and this means that we must assume that $y \neq 0$ (i.e., y is not the constant function equal to 0). Remember this moment – we will return to it at the end of the example.

$$\frac{dy}{y} = \frac{dx}{7x^2},$$

$$\int \frac{dy}{y} = \int \frac{dx}{7x^2},$$

$$\ln|y| = \frac{-1}{7x} + C .$$

This time it was impossible to immediately obtain the formula for y, so we have to transform the expression. We will use the fact that the inverse of a logarithmic function is the exponential function.

$$e^{\ln|y|} = e^{\frac{-1}{7x}+C},$$

$$|y| = e^{\frac{-1}{7x}} \cdot e^{C} .$$

We can remove the absolute value:

$$y = e^{\frac{-1}{7x}} \cdot e^{C} \quad \vee \quad y = -e^{\frac{-1}{7x}} \cdot e^{C} .$$

The right hand side of the first equation ends with e^C. It is a constant. As C can be arbitrary chosen, e^C is a positive number. Since after removing the absolute value we have to consider also the second equation, we can write the solution in the form:

$$y = C_1 e^{\frac{-1}{7x}} .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Here $C_1 = \pm C$, hence $C_1 \in \mathrm{R} \setminus \{0\}$. Now recall the beginning of this example. We excluded the function $y = 0$. We still have to check, whether it satisfies the initial equation. Assume that $y = 0$ is the constant function equal to 0. Taking into consideration that the derivative of y (i.e. dy / dx) would be also 0 in this case, the left side of the equation takes the form:

$$L = 7x^2 \frac{dy}{dx} = 7x^2 \cdot 0 = 0 .$$

As the right side is $R = y = 0$, we obtain $L = R$, hence the function $y = 0$ satisfies the equation, too. It follows that we can extend the domain of the constant to all the real numbers:

$$y = C_1 e^{\frac{-1}{7x}}, \quad C_1 \in R .$$

In such situation we do not write the assumptions. Last, cosmetic change involves the removal of the index at constant (indices appear in each step of the solution in order to distinguish the constants, but finally we are only interested in the information on where the constant is located). So the solution looks as follows:

$$y = Ce^{\frac{-1}{7x}} .$$

Example 3
Solve the differential equation:

$$\frac{dy}{dx} = 3x - 7y + 11 .$$

Solution
We will solve the equation by constant coefficient substitution. We introduce new variable:

$$u = 3x - 7y + 11 .$$

Hence

$$y = \frac{3x - u + 11}{7},$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\frac{dy}{dx} = -\frac{1}{7} \cdot \frac{du}{dx} + \frac{3}{7}.$$

The equation takes the form:

$$-\frac{1}{7} \cdot \frac{du}{dx} + \frac{3}{7} = u.$$

After some basic transformations we obtain:

$$\frac{du}{dx} = -7u + 3.$$

And finally (remember the assumption $u \neq 3/7$, which has to be verified later):

$$\frac{du}{u - 3/7} = -7dx.$$

The general integral of this equation (after taking into account the case $u = 3/7$) has the form:

$$u = Ce^{-7x} + 3/7.$$

Finally we go back to the previous notation. Since $y = (3x - u + 11)/7$, the solution of the initial equation is:

$$y = \frac{3x - Ce^{-7x} + 74/7}{7}.$$

Example 4
Solve the differential equation:

$$x\frac{dy}{dx} = x + y.$$

Solution
This time we are dealing with a homogeneous equation (after dividing both sides by x we obtain the form

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\frac{dy}{dx} = 1 + \frac{y}{x},$$

in which x and y are present only together as the fraction y / x). We use the appropriate substitution:

$$u = \frac{y}{x}.$$

Hence

$$y = ux,$$

$$\frac{dy}{dx} = \frac{du}{dx}x + u.$$

The equation takes the form:

$$x\left(\frac{du}{dx}x + u\right) = x + ux.$$

After further reduction we obtain:

$$\frac{du}{dx}x + u = 1 + u,$$

$$\frac{du}{dx}x = 1,$$

$$du = \frac{dx}{x}.$$

The general integral is:

$$u = \ln|x| + C.$$

Going back to the previous notation we obtain $y = ux$, i.e.

$$y = x(\ln|x| + C).$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 5
Solve the initial value problem:

$$\begin{cases} \dfrac{dy}{dx} = 12x, \\ f(2) = 33. \end{cases}$$

Solution
We begin with solving the differential equation. It is an equation with separated variables, so the solution is as follows:

$dy = 12x\,dx$,

$\int dy = \int 12x\,dx$,

$y = 6x^2 + C$.

Knowing the general integral of the equation, we use the initial condition to find the value of C:

$f(2) = 33$,

$6 \cdot 2^2 + C = 33$.

Hence $C = 9$. Finally, the solution of the problem is:

$y = 6x^2 + 9$.

EXERCISES

1. Solve the differential equations:

a) $\dfrac{dy}{dx} = 2x + 7$;

b) $\dfrac{dy}{dx} = xe^x$;

c) $\dfrac{dy}{dx} = yx$;

d) $2x^3 \dfrac{dy}{dx} = y$;

e) $x\dfrac{dy}{dx} = y^2$;

f) $x\dfrac{dy}{dx} = x + 7y$;

g) $x\dfrac{dy}{dx} = 7x + y$;

h) $\dfrac{dy}{dx} = 7x + y + 11$;

i) $\dfrac{dy}{dx} = 11x + y + 7$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Solve the initial value problems:

a) $\begin{cases} \dfrac{dy}{dx} = \ln x, \\ f(1) = 5; \end{cases}$ b) $\begin{cases} \dfrac{dy}{dx} = yx^2, \\ f(0) = 7; \end{cases}$ c) $\begin{cases} x\dfrac{dy}{dx} = -y^2, \\ f(-1) = 1. \end{cases}$

SOLUTIONS

1. a) $y = x^2 + 7x + C$; b) $y = xe^x - e^x + C$;

c) $y = Ce^{1/2x^2}$; d) $y = Ce^{-\frac{1}{4x^2}}$;

e) $y = \dfrac{-1}{\ln|x| + C}$ or $y = 0$;

f) (homogeneous equation) $y = x(Cx^6 - 1/6)$;

g) (homogeneous equation) $y = x(7\ln|x| + C)$;

h) $y = Ce^x - 7x - 18$; i) $y = Ce^x - 11x - 18$.

2. a) $y = x(\ln|x| - 1) + 5$; b) $y = 7e^{1/3x^3}$; c) $y = \dfrac{1}{\ln|x| + 1}$.

11.2. DIFFERENCE EQUATIONS

THEORY IN A NUTSHELL

As you already know, differential equations describe certain relations between functions and their derivatives. The equivalent to the differential equations for sequences are difference equations.

Definition 11.8. Difference equation[26]
An equation of the form:

$F(n, a_n, a_{n-1}, a_{n-2}, \ldots, a_{n-k}) = 0,$

where the unknown is sequence (a_n), is called ***k*–th order difference equation**.

In the remainder, we will be interested in linear homogeneous difference equations with constant coefficients.

[26] The equations of this kind are also called **recurrence relations**. Some authors define the difference equations as a special kind of recurrence relations. In this book we will apply another approach, where these two notions are considered as equivalent.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Definition 11.9. Linear homogeneous difference equation with constant coefficients
Linear homogeneous difference equation with constant coefficients is an equation of the form:

$$a_n + p_1 a_{n-1} + p_2 a_{n-2} + \ldots + p_k a_{n-k} = 0.$$

Due to the complexity, we will focus on first and second order equations. The following theorem presents methods of solving this type of equations.

Theorem 11.1. Solutions of the first and second order linear homogeneous difference equations with constant coefficients

1. The solution of an equation of order 1 of the form $a_n + p_1 a_{n-1} = 0$ is described with the formula $a_n = C(-p_1)^n$.
2. The solution of an equation of order 2 of the form $a_n + p_1 a_{n-1} + p_2 a_{n-2} = 0$ is described with the formula:
   a) $a_n = C_1 x_1^n + C_2 x_2^n$, if the quadratic equation $x^2 + p_1 x + p_2 = 0$ has two distinct real solutions x_1 and x_2;
   b) $a_n = x_0^n (C_1 n + C_2)$, if the quadratic equation $x^2 + p_1 x + p_2 = 0$ has one double real solution x_0;
   c) $a_n = \left(C_1 \sin(n\varphi) + C_2 \cos(n\varphi)\right)\left(\sqrt{p_2}\right)^n$, if the quadratic equation $x^2 + p_1 x + p_2 = 0$ has no real solutions, i.e. the discriminant $\Delta = p_1^2 - 4p_2 < 0$; here $\varphi \in (0, 2\pi)$ is angle such that $\sin\varphi = \frac{\sqrt{-\Delta}}{2\sqrt{p_2}}$ and $\cos\varphi = \frac{-p_1}{2\sqrt{p_2}}$.

Similarly as in the case of differential equations, we can define the **initial value problem** for a difference equation. For an equation of order k it consists of the equation and k first terms of the searched sequence.
If you want to solve a difference equation or initial value problem using **WolframAlpha®**, just type it. For example, in order to solve Example 2 below, you should type:

"a(n)-4a(n-1)+3a(n-2)=0".

EXAMPLES

Example 1
Solve the difference equation:

$$a_n + 4a_{n-1} = 0.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
It is a first order linear homogeneous difference equation with constant coefficients. Using the formula (property 1), we obtain the solution:

$$a_n = C(-4)^n .$$

Example 2
Solve the difference equation:

$$a_n - 4a_{n-1} + 3a_{n-2} = 0 .$$

Solution
This time we deal with a second order linear homogeneous difference equation with constant coefficients. The characteristic equation is

$$x^2 - 4x + 3 = 0 .$$

The solutions of this equation are $x = 1$ and $x = 3$. Hence the solution of the difference equation is:

$$a_n = C_1 \cdot 1^n + C_2 \cdot 3^n = C_1 + C_2 \cdot 3^n .$$

Example 3
Solve the difference equation:

$$a_n - 4a_{n-1} + 4a_{n-2} = 0 .$$

Solution
This time we also deal with a second order linear homogeneous difference equation with constant coefficients. The characteristic equation is

$$x^2 - 4x + 4 = 0 .$$

It has one double root: $x = 2$. Hence its solution is the following family of sequences:

$$a_n = 2^n (C_1 n + C_2) .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 4
Solve the difference equation:

$$a_n = 2a_{n-1} - 4a_{n-2}.$$

Solution
Once again we deal with a second order linear homogeneous difference equation with constant coefficients. The characteristic equation is

$$x^2 - 2x + 4 = 0.$$

It has negative discriminant:

$$\Delta = -12 < 0.$$

We find the angle φ by solving the system of equations:

$$\begin{cases} \sin\varphi = \dfrac{\sqrt{12}}{2\sqrt{4}} = \dfrac{\sqrt{3}}{2}, \\ \cos\varphi = \dfrac{2}{2\sqrt{4}} = \dfrac{1}{2}. \end{cases}$$

It follows that $\varphi = \dfrac{\pi}{3}$, so the solution of the difference equation is:

$$a_n = \left(C_1 \sin\left(\frac{n\pi}{3}\right) + C_2 \cos\left(\frac{n\pi}{3}\right)\right) \cdot 2^n.$$

Example 5
Assume that a sequence is defined with the following conditions[27]: $a_1 = 1$, $a_2 = 1$, $a_n = a_{n-1} + a_{n-2}$ for $n = 3, 4, 5, \ldots$ Find te formula for the n-th term of this sequence.

Solution
The difference equation in the initial value problem under discussion is $a_n - a_{n-1} - a_{n-2} = 0$. Thus the characteristic equation is

$$x^2 - x - 1 = 0,$$

[27] These conditions define the so-called Fibonacci sequence.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

and its solutions are $x_1 = \frac{1-\sqrt{5}}{2}$ and $x_2 = \frac{1+\sqrt{5}}{2}$. It follows that its solution is the family of sequences:

$$a_n = C_1\left(\frac{1-\sqrt{5}}{2}\right)^n + C_2\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

We substitute the initial conditions $a_1 = 1$ and $a_2 = 1$:

$$a_1 = C_1\left(\frac{1-\sqrt{5}}{2}\right)^1 + C_2\left(\frac{1+\sqrt{5}}{2}\right)^1 = 1,$$

$$a_2 = C_1\left(\frac{1-\sqrt{5}}{2}\right)^2 + C_2\left(\frac{1+\sqrt{5}}{2}\right)^2 = 1.$$

The solution of the above linear system is $C_1 = -1/\sqrt{5}$ and $C_2 = 1/\sqrt{5}$. It follows that the solution of the problem is described with the formula:

$$a_n = -\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Note that all the terms of the resulting sequence are natural numbers (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, …), although it is hard to believe, when looking at the last formula.

EXERCISES

1. Solve the difference equations:
   a) $a_n + 5a_{n-1} = 0$;
   b) $a_n - 7a_{n-1} = 0$;
   c) $a_n - 5a_{n-1} + 6a_{n-2} = 0$;
   d) $a_n - 6a_{n-1} + 8a_{n-2} = 0$;
   e) $a_n - 6a_{n-1} + 9a_{n-2} = 0$;
   f) $a_n - 10a_{n-1} + 25a_{n-2} = 0$.
   g) $a_n - 4\sqrt{2}a_{n-1} + 16a_{n-2} = 0$;
   f) $a_n + 4\sqrt{2}a_{n-1} + 16a_{n-2} = 0$.
2. Solve the initial value problems:
   a) $\begin{cases} a_n - 11a_{n-1} = 0, \\ a_1 = 22; \end{cases}$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

b) $\begin{cases} a_n - 9a_{n-1} + 20a_{n-2} = 0, \\ a_0 = 6, \\ a_1 = 28; \end{cases}$

c) $\begin{cases} a_n + 12a_{n-1} + 36a_{n-2} = 0, \\ a_1 = -30, \\ a_2 = 288. \end{cases}$

SOLUTIONS

1. a) $a_n = C(-5)^n$; b) $a_n = C \cdot 7^n$;

   c) $a_n = C_1 \cdot 2^n + C_2 \cdot 3^n$; d) $a_n = C_1 \cdot 2^n + C_2 \cdot 4^n$;

   e) $a_n = 3^n (C_1 n + C_2)$; f) $a_n = 5^n (C_1 n + C_2)$;

   g) $a_n = \left(C_1 \sin\left(\frac{n\pi}{4}\right) + C_2 \cos\left(\frac{n\pi}{4}\right)\right) \cdot 4^n$;

   h) $a_n = \left(C_1 \sin\left(\frac{3n\pi}{4}\right) + C_2 \cos\left(\frac{3n\pi}{4}\right)\right) \cdot 4^n$.

2. a) $a_n = 2 \cdot 11^n$; b) $a_n = 2 \cdot 4^n + 4 \cdot 5^n$;

   c) $a_n = (-6)^n (3n + 2)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 12
APPLICATIONS OF CALCULUS IN ECONOMICS

12.1. DIFFERENTIAL CALCULUS

THEORY IN A NUTSHELL

Many economic quantities can be expressed with functions. For example, the cost of production depends on its size (in this case we are talking about the cost function), revenue on sales volume (a function of revenue), the demand on price and consumers' income (demand function), sales changes over time (sales trend), etc. The concepts introduced in this chapter allow analysis of all (or almost all) of these functions.

Definition 12.1. Unit value
Unit value of a function is the part of its value per one unit of the argument, described in units. In the case of a function of one variable, it is described with the formula:

$$U_f(x) = \frac{f(x)}{x},$$

while in the case of a function of many variables, with the formula:

$$U_f^{x_j}(x) = \frac{f(x_1, x_2, ..., x_n)}{x_j}$$

(in this case it can be derived independently for each argument).

Definition 12.2. Marginal value
Marginal value is the change of the value of a function (in units) corresponding with the change of variable by one unit. In the case of a function of one variable, it is described with the formula:

$$P_f(x) = f'(x),$$

while in the case of a function of many variables, with the formula:

$$P_f^{x_j}(x) = f'_{x_j}(x_1, x_2, ..., x_n)$$

(in this case it can be derived independently for each argument).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Definition 12.3. Growth rate
Growth rate is the change of the value of a function (in percents) corresponding with the change of variable by one unit (you need to multiply the result by 100%). In the case of a function of one variable, it is described with the formula:

$$S_f(x) = \frac{f'(x)}{f(x)},$$

while in the case of a function of many variables, with the formula:

$$S_f^{x_j}(x) = \frac{f'_{x_j}(x_1, x_2, ..., x_n)}{f(x_1, x_2, ..., x_n)}$$

(in this case it can be derived independently for each argument).

Definition 12.4. Elasticity
Elasticity is the change of the value of a function (in percents) corresponding with the change of variable by one percent. In the case of a function of one variable, it is described with the formula:

$$E_f(x) = \frac{f'(x)}{f(x)} \cdot x,$$

while in the case of a function of many variables, with the formula:

$$E_f^{x_j}(x) = \frac{f'_{x_j}(x_1, x_2, ..., x_n)}{f(x_1, x_2, ..., x_n)} \cdot x_j$$

(in this case it can be derived independently for each argument).

EXAMPLES

Example 1
The company's production costs are described with the formula (x – production in thousand pieces, K – total cost in million PLN): $K(x) = x^2 + 4x + 16$. Find the volume of production that minimizes the unit cost. Find and interprete the elasticity of total cost for this volume of production.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
First, we find the formula of the unit cost:

$$U_K(x) = \frac{K(x)}{x} = \frac{x^2 + 4x + 16}{x} = x + 4 + \frac{16}{x}.$$

Domain of this function is $D(U_K) = R^+$ (of course this function is defined also for $x < 0$, but you must remember that the production volume cannot be negative). In order to find the extrema, we have to derive its derivative:

$$U_K'(x) = 1 - \frac{16}{x^2}.$$

The domain of derivative is $D(U_K') = R^+$. It means that there are no type 2 stationary points. In order to find the type 1 stationary points, we find the zeros of the derivative:

$$U_K'(x) = 0 \quad \Leftrightarrow \quad 1 - \frac{16}{x^2} = 0 \quad \Leftrightarrow \quad x = 4 \vee x = -4.$$

Taking into account the assumptions, we reject the negative value, hence $x = 4$. In order to verify whether it is maximum or minimum, we determine the intervals of increase and decrease:

$$U_K'(x) > 0 \quad \Leftrightarrow \quad 1 - \frac{16}{x^2} > 0 \quad \Leftrightarrow \quad x \in (-\infty, -4) \cup (4, \infty),$$

$$U_K'(x) < 0 \quad \Leftrightarrow \quad 1 - \frac{16}{x^2} < 0 \quad \Leftrightarrow \quad x \in (-4, 0) \cup (0, 4).$$

Bearing in mind the assumption $x > 0$, we finally conclude that the function decreases when $x \in (0, 4)$ and increases when $x \in (4, \infty)$. This means that it has a minimum at $x = 4$. Thus the company minimizes its unit cost, when producing 4,000 pieces. We find the elasticity using the formula:

$$E_K(x) = \frac{K'(x)}{K(x)} \cdot x = \frac{2x + 4}{x^2 + 4x + 16} \cdot x = \frac{2x^2 + 4x}{x^2 + 4x + 16}.$$

Its value at $x = 4$ is:

$$E_K(4) = \frac{2 \cdot 4^2 + 4 \cdot 4}{4^2 + 4 \cdot 4 + 16} = 1.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

It means that for the production volume equal to 4,000 pieces, an increase of production by 1% would cause the increase of total cost by about 1%.

Example 2
The demand for company's products in years 2004–2013 is described with the formula: (D – demand in thousand pieces, t – time in years, $t = 0$ for the year 2003): $D(t) = 100 \exp(0.03t)$. Find and interprete the growth rate of the demand function.

Solution
Let us recall that $\exp(x)$ is another notation for e^x. In order to find the growth rate, we use the formula:

$$S_D(t) = \frac{D'(t)}{D(t)} = \frac{100\exp(0.03t) \cdot 0.03}{100\exp(0.03t)} = 0.03 .$$

This value means that from year to year (i.e., when t increases by 1) the demand increases by about 3%. Note that the growth rate does not depend on t (i.e., it is constant). As you will see soon, exponential functions are the only functions having this property.

Example 3
The company's unit production cost is described with the formula:

$$KP(x) = 60 + 0.01x + \frac{150}{x},$$

and the unit price with the formula:

$$p(x) = 120 - 0.02x .$$

Find the production volume x maximizing the profit.

Solution
Let us denote by x the volume of production, by K the function of total cost, and by P – the function of total income. It follows that

$$U_K(x) = 60 + 0.01x + \frac{150}{x},$$

so we can derive the function of total cost:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\frac{K(x)}{x} = 60 + 0.01x + \frac{150}{x},$$

$$K(x) = \left(60 + 0.01x + \frac{150}{x}\right) \cdot x = 0.01x^2 + 60x + 150.$$

To get the revenue function, we have to multiply the price by the volume of sales (assuming that it is equal to the volume of production):

$$P(x) = (120 - 0.02x) \cdot x = -0.02x^2 + 120x.$$

Then we can determine the function of profit (as the difference between revenue and cost):

$$Z(x) = P(x) - K(x) = (-0.02x^2 + 120x) - (0.01x^2 + 60x + 150) =$$
$$= -0.03x^2 + 60x - 150.$$

Now we have to find the maximum. The derivative is:

$$Z'(x) = -0.06x + 60.$$

As you can easily check:

$$Z'(x) = 0 \quad \Leftrightarrow \quad x = 1{,}000,$$

$$Z'(x) > 0 \quad \Leftrightarrow \quad x < 1{,}000,$$

$$Z'(x) < 0 \quad \Leftrightarrow \quad x > 1{,}000.$$

It means that the company maximizes its profit when the production volume is 1000.

Example 4
The demand function is described with the formula:

$$D(p) = 2{,}500 + 10p - 0.3p^2,$$

and the supply function with the formula:

$$S(p) = 2{,}100 + 4p + 0.1p^2$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

(S – supply in thousand pieces, D – demand in thousand pieces, p – price in PLN). Find the equilibrium price, and then find and interprete the price elasticity of demand and supply for this price.

Solution
In order to find the equilibrium price, we need to equate the demand and the supply:

$$2{,}100 + 4p + 0.1p^2 = 2{,}500 + 10p - 0.3p^2 .$$

Hence $0.4p^2 - 6p - 400 = 0$, i.e. $p = -25 \vee p = 40$. As we have to reject the negative value, the equilibrium price is 40 PLN. The elasticity of demand is described with the formula:

$$E_D(p) = \frac{10p - 0.6p^2}{2{,}500 + 10p - 0.3p^2} .$$

The elasticity of supply is in turn equal to:

$$E_S(p) = \frac{4p + 0.2p^2}{2{,}100 + 4p + 0.1p^2} .$$

Their values at $p = 40$ are respectively $E_D(40) = -560/2{,}420 \approx -0.231$ and $E_S(40) = 480/2{,}420 \approx 0.198$. It means that at the equilibrium point, the increase of price by 1% would cause decrease of demand by about 0.231% and increase of supply by about 0.198%.

Example 5
The changes of demand over time are described with the function:

$$D(t) = 10{,}000\left(a - \frac{1}{t+b}\right),$$

where D – demand in thousand pieces, t – time in years. It is known that in the beginning of the analyzed period the demand was equal to 0 and after five years it increased by 40/3% yearly. What is the exact formula of the function (find a and b)? What is the demand forecast for the end of the tenth year of observations?

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
Year to year 40/3% growth means that the growth rate at $t = 5$ equals to 40/3 / 100 = 2/15. On the other hand, the value of demand at $t = 0$ was 0. We can write these conditions in the form of the following system of equations:

$$\begin{cases} D(0) = 0, \\ S_D(5) = 2/15. \end{cases}$$

As the growth rate of D is described with the formula

$$S_D(t) = \frac{10{,}000\left(\frac{1}{t+b}\right)^2}{10{,}000\left(a - \frac{1}{t+b}\right)} = \frac{1}{(t+b)[a(t+b)-1]},$$

this system takes the form:

$$\begin{cases} 10{,}000\left(a - \frac{1}{0+b}\right) = 0, \\ \frac{1}{(5+b)[a(5+b)-1]} = 2/15. \end{cases}$$

From the first equation it follows that $a = 1/b$. By substituting it to the second equation we obtain: $(5 + b)(5/b + 1 – 1) = 15/2$, i.e. $b = 10$, and hence $a = 0.1$. Finally the formula of the function is:

$$D(t) = 10{,}000\left(0.1 - \frac{1}{t+10}\right).$$

Since $D(10) = 500$, one should expect the demand equal to 500,000 pieces in the end of the tenth year of observations.

Example 6
Determine the dimensions of a cuboid box having volume of 1 dm^3 (1 liter) in such a way that the production cost will be minimum.

Solution
Let us denote the dimensions by x, y i z. The volume equals to:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$V(x, y, z) = xyz = 1.$

The production cost is directly proportional to the amount of material used, or the surface of the cuboid:

$S(x, y, z) = 2xy + 2xz + 2yz.$

It follows that the problem is to find the minimum of the function $S(x, y, z) = 2xy + 2xz + 2yz$ subject to $xyz - 1 = 0$. The Lagrange function is:

$$L(u, x, y, z) = 2xy + 2xz + 2yz + u(xyz - 1).$$

Its gradient is:

$$\nabla L(u, x, y, z) = [xyz - 1, \quad 2y + 2z + uyz, \quad 2x + 2z + uxz, \quad 2x + 2y + uxy]^T.$$

To find the stationary points, we solve the system of equations:

$$\begin{cases} xyz = 1, \\ 2y + 2z = -uyz, \\ 2x + 2z = -uxz, \\ 2x + 2y = -uxy. \end{cases}$$

By dividing the second equation by the third one (remember about the assumptions!), we obtain

$$\frac{y + z}{x + z} = \frac{y}{x},$$

hence $xy + xz = xy + yz$, i.e. $xz = yz$, so $x = y$ (from the assumptions it follows that $z \neq 0$). Similarly, after dividing the second equation by the fourth one we obtain $x = z$, thus finally $x = y = z = 1$, $u = -4$. The Hessian matrix of the Lagrange function is:

$$HL(u, x, y, z) = \begin{bmatrix} 0 & yz & xz & xy \\ yz & 0 & 2 + uz & 2 + uy \\ xz & 2 + uz & 0 & 2 + ux \\ xy & 2 + uy & 2 + ux & 0 \end{bmatrix},$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$HL(-4,1,1,1)=\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & -2 & -2 \\ 1 & -2 & 0 & -2 \\ 1 & -2 & -2 & 0 \end{bmatrix}.$$

By calculating the relevant determinants, we get confirmation that there is a minimum at the analyzed point. So, to minimize the cost of manufacturing the box, one should give it the shape of a cube with edge length of 1 dm (10 cm).

EXERCISES

1. The company's total revenue is defined by the function $P(x) = -10x^2+200x$, and the total cost by the function $K(x) = 10x^2 + 100x + 40$, where x – production in tonnes, P – revenue in thousand PLN, K – cost in thousand PLN.
    a) How much will the profit of the company for the production equal to 5 tons?
    b) Determine the volume of production that maximizes profit and find the maximum value of profit.
2. The company's total revenue is defined by the function $P(x) = -5x^2 + 80x$, and the total cost by the function $K(x) = 5x^2 + 30x + 100$, where x – production in tonnes, P – revenue in thousand PLN, K – cost in thousand PLN.
    a) How much will the profit of the company for the production equal to 3 tons?
    b) Determine the volume of production that maximizes profit and find the maximum value of profit.
    c) What the management of the company is supposed to do?
3. The company's total revenue is defined by the function $P(x) = -5x^2 + 200x$, and the total cost by the function $K(x) = 10x^2 + 80x + 10$, where x – production in tonnes, P – revenue in thousand PLN, K – cost in thousand PLN.
    a) Determine the volume of production that maximizes profit and find the maximum value of profit.
    b) Find the elasticity of total cost for the optimal volume of production.
4. The company's total revenue is defined by the function $P(x) = -3x^2 + 200x$, and the total cost by the function $K(x) = 9x^2 + 80x + 36$, where x – production in tonnes, P – revenue in thousand PLN, K – cost in thousand PLN.
    a) Determine the volume of production that maximizes profit and find the maximum value of profit.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

b) Determine the volume of production that minimizes the unit cost and find the minimum value of that cost.

5. The company "Cheap Train" wants to analyze the demand on its services. It follows from the research that the sales should be described by the exponential function $y = 27e^{0.06t}$, where t – time in years ($t = 0$ for the year 2005), and y – annual ticket sales in millions. What is the approximate year to year increase of sales and what is the forecast for the year 2016?
6. Company manufactures two products X and Y. The total production cost in million PLN is described by the function: $f(x, y) = 1{,}000\ln(2x^2 + y^2)$, where x and y are the production volume of X and Y, respectively (in thousand pieces). During the production process, one resource is used: 6 kg per one piece of X and 3 kg per one piece of Y. Its weekly limit equals to 90 tons and must be fully used. Find the optimum plan of production.
7. What shape should have an open cuboid box having the volume of 13.5 l to make its production cost minimum?
8. What shape should have an open cuboid box having surface of 75 dm^2 to make its volume maximum?
9. What shape should have a closed cuboid box having surface of 96 dm^2 to make its volume maximum?

SOLUTIONS

1. a) $Z(5) = -40$; b) $Z_{max}(2.5) = 85$.
2. a) $Z(3) = -40$; b) $Z_{max}(2.5) = -37.5$;
   c) It is necessary to reduce the fixed costs. If it does not help, the company should be closed.
3. a) $Z_{max}(4) = 230$; b) $E_K(4) = 64/49$.
4. a) $Z_{max}(5) = 290$; b) $U_{K\min}(2) = 116$.
5. About 6%; In 2016 it will be approximately 52.24 mln.
6. Problem: $f(x, y) = 1{,}000\ln(2x^2 + y^2) \to \min$ subject to $6x + 3y = 90$; solution: $x = y = 10$.
7. Dimensions of the base: x, y, height z; problem: $S(x) = xy + 2xz + 2yz \to \min$ subject to $V(x, y, z) = xyz = 13.5$; solution: $x = 3$, $y = 3$, $z = 1.5$.
8. Dimensions of the base: x, y, height z; problem: $V(x, y, z) = xyz \to \max$ subject to $S(x) = xy + 2xz + 2yz = 75$; solution: $x = 5$, $y = 5$, $z = 2.5$.
9. Dimensions of the base: x, y, height z; problem: $V(x, y, z) = xyz \to \max$ subject to $S(x) = 2xy + 2xz + 2yz = 96$; solution: $x = 4$, $y = 4$, $z = 4$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

12.2. IMPORTANT FUNCTIONS

THEORY IN A NUTSHELL

Linear function $f(x) = b_1x + b_0$ (in case of many variables $f(x_1, x_2, \ldots, x_n) = b_1x_1 + b_1x_2 + \ldots + b_1x_n + b_0$) is the only function having **constant marginal value**, equal to b_1 (in case of many variables equal to b_j with respect to x_j). Each time when you observe that the change of the value of a function (in units) corresponding with the change of variable by one unit is constant, it means that the relation is described with a linear function. Functions of this kind are used in modelling the changes of some phenomenons in time, as well as the short-term relations between production and costs.

Exponential function $f(x) = Ae^{bx}$ (in case of many variables $f(x) = Ae^{b_1x_1+b_2x_2+\ldots+b_nx_n}$) is the only function having **constant growth rate**, equal to b (in case of many variables equal to b_j with respect to x_j). Each time when you observe that the change of the value of a function (in percents) corresponding with the change of variable by one unit is constant, it means that the relation is described with an exponential function. Functions of this kind are used in modelling the changes in time of prices, population or GDP.

Power function (Cobb–Douglas function) $f(x) = Ax^b$ (in case of many variables $f(x) = Ax_1^{b_1}x_2^{b_2}\ldots x_n^{b_n}$) is the only function having **constant elasticity**, equal to b (in case of many variables equal to b_j with respect to x_j). Each time when you observe that the change of the value of a function (in percents) corresponding with the change of variable by one percent is constant, it means that the relation is described with a power function. Functions of this kind are used in modelling the relations between demand, prices and incomes or production and the usage of resources.

Logistic function $f(x) = \dfrac{a}{1+be^{-cx}}$ is an increasing function such that first its growth increases (until the function gets the value $a/2$), and then the growth slows, so that the function approaches certain constant level a, called **saturation level**. It has the following property: at each point its growth is directly proportional to both its current value and the difference between the saturation level and the current value. Sample shape of the logistic function has been presented in Fig. 12.1a. Functions of this kind are used in modelling the daily (weekly, monthly) changes in sales of new, innovative products.

Type 1 Törnquist function (function T1) $f(x) = \dfrac{ax}{x+b}$ is increasing and its growth is decreasing, while the value of f approaches the saturation level a. Sample shape of function T1 has been presented in Fig. 12.1b. Functions of this

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

kind reflect the dependence between the demand for necessity goods and the consumer's income.

Type 2 Törnquist function (function T2) $f(x) = \dfrac{a(x-c)}{x+b}$ is increasing and its growth is decreasing, while the value of f approaches the saturation level a. It is defined only for $x \geq c$. Sample shape of function T2 has been presented in Fig. 12.1c. Functions of this kind reflect the dependence between the demand for superior goods and the consumer's income.

Type 3 Törnquist function (function T3) $f(x) = \dfrac{ax(x-c)}{x+b}$ is increasing and its growth is increasing, while the value of f tends to infinity. It is defined only for $x \geq c$. Sample shape of function T3 has been presented in Fig. 12.1d. Functions of this kind reflect the dependence between the demand for luxury goods and the consumer's income.

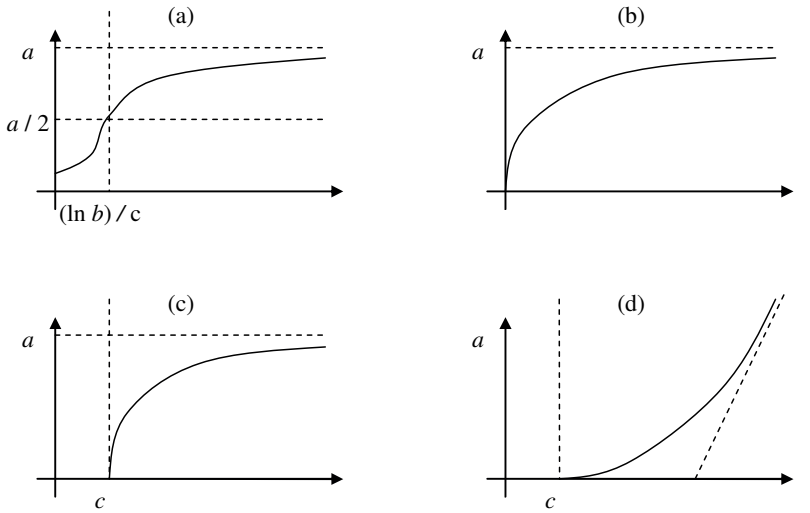


Fig. 12.1. Logistic function and Törnquist functions

EXAMPLES

Example 1

Find the formula of the function having constant growth rate equal to b.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
Let us denote the function by y and its argument by x. We have to solve the equation $S_y(x) = b$. Bearing in mind that the derivative can be denoted by dy / dx, the formula for the growth rate can be written as:

$$S_y(x) = \frac{dy/dx}{y} \text{ (or: } S_y(x) = \frac{1}{y} \cdot \frac{dy}{dx}\text{).}$$

It means that we are going to solve the following differential equation:

$$\frac{1}{y} \cdot \frac{dy}{dx} = b\,.$$

After standard transformations we obtain:

$$\frac{dy}{y} = b\,dx\,,$$

$$\int \frac{dy}{y} = \int b\,dx\,.$$

Thus finally:

$$y = Ce^{bx}, \quad C \neq 0\,.$$

We are not able to find the exact formula of the function, since we do not know the initial conditions. However, as we can see, it is for sure an exponential function.

Example 2
The number of new customers of new mobile network provider is changing in such a way that its growth is directly proportional to both the number of existing subscribers of the network and the difference between this number and a hypothetical saturation level a. What model describes the variation of subscribers in time?

Solution
Let us denote time by x, and the number of subscribers by y. The growth of the number of subscribers is equal to the derivative dy / dx, which is directly proportional to y and to $a - y$. It means that the function must satisfy the following differential equation (p denotes the coefficient of proportionality):

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\frac{dy}{dx} = py(a-y)\,.$$

We transform the equation and integrate both sides (we can exclude the cases $y = 0$ and $y = a$ as they both mean lack of dynamics):

$$\int \frac{dy}{y(a-y)} = \int p dx\,.$$

In order to integrate the left hand side, note that:

$$\frac{1}{y} + \frac{1}{a-y} = \frac{a}{y(a-y)},$$

hence:

$$\frac{1}{y(a-y)} = \frac{1}{a}\left(\frac{1}{y} + \frac{1}{a-y}\right)\,.$$

Thus we obtain:

$$\frac{1}{a}\int\left(\frac{1}{y} + \frac{1}{a-y}\right)dy = \int p\,dx\,,$$

$$\frac{1}{a}(\ln|y| - \ln|y-a|) = px + C\,.$$

Now we transform the expressions using, among others, the properties of operations on logarithms:

$$-\frac{1}{a}\ln\left|\frac{y-a}{y}\right| = px + C\,,$$

$$\ln\left|1 - \frac{a}{y}\right| = -apx - aC\,.$$

Since $y < a$ (y does not reach the saturation level), we can write it in the form:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$\ln\left(\frac{a}{y}-1\right)=-apx-aC\,.$$

Hence:

$$\frac{a}{y}-1=e^{-apx-aC}=e^{-apx}\cdot e^{-aC}\,.$$

Finally we obtain:

$$y=\frac{a}{1+e^{-apx}\cdot e^{-aC}}\,.$$

Denoting e^{-aC} by b and $-ap$ by c, we obtain the formula for the family of functions satisfying the given conditions:

$$y=\frac{a}{1+be^{-cx}}\,.$$

As we can see, the number of subscribers is described by the logistic function.

Example 3
Find the saturation level of the logistic function:

$$y=\frac{1{,}300}{1+2e^{-0{,}04x}}\,.$$

Solution
In fact, you already know how much this level is, so the task is rather to confirm your knowledge. The saturation level is the level to which approaches the value of the function when the argument tends to infinity. So we need to determine the appropriate limit:

$$\lim_{x\to\infty}\frac{1{,}300}{1+2e^{-0{,}04x}}=\left[\frac{1{,}300}{1+2e^{-\infty}}\right]=\frac{1{,}300}{1+2\cdot 0}=1{,}300\,.$$

Example 4
Demand for good D (in thousand pieces) depends on its price P (PLN) and advertising expenses R (thousand PLN). It is known that 1% increase of price is followed by 1% decrease of demand, while 1% increase of advertising expenses

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

causes 0.5% increase of demand. When the price is 10 PLN and the advertising expenses are 25 thousand PLN, the demand equals to 100 thousand pieces. Find the demand function. What should be the price if the advertising expenses are 100 thousand PLN and we want the demand to be 400 thousand pieces?

Solution

Taking into account the constraints, in a general case we should solve the relevant initial value problem. However, since the function has constant elasticities (–1 with respect to the price and 0.5 with respect to the advertising expenses), we can be sure that it is a Cobb–Douglas function. Its formula is: $D(P, R) = CP^{-1}R^{0.5}$, where C is a constant. In order to find it we substitute the initial condition $D(10, 25) = 100$: $C \cdot 10^{-1} \cdot 25^{0.5} = 100$. Hence $C = 200$. Finally, the demand function is $D(P, R) = 200P^{-1}R^{0.5}$. In order to find the desired price, we substitute the demand and advertising expenses: $D(P, 100) = 400$, i.e. $200P^{-1}100^{0.5} = 400$, thus $P = 5$. The price should be equal to 5 PLN.

Example 5

The company has established the following relationship between the demand for margarine, measured in terms of sales S (tons per week), and the average income of consumers D (in PLN for 1 person per month) and the price of margarine C (in PLN per 1 kg): $S = 20\, D^{0.5}\, C^{-1.2}$. Also the relationship between the weekly production cost K (thousand PLN) and production volume Q (in tons per week) has been estimated: $K = 100 + 3Q$. We assume that the weekly production is adapted to the volume of sales. Current price of margarine $C_1 = 6$ PLN/kg, and the sales volume $S_1 = 50$ tons. Calculate the approximate change of demand for margarine, if its price falls by 10%. Determine how in this case will change weekly revenue, cost and profit.

Solution

Since power function describes the sales, its price elasticity is constant (equal to –1.2). It means that 10% decrease of price, will cause the demand decrease by about $1.2\% \cdot 10 = 12\%$. Thus the new price will be $6 - 10\% \cdot 6 = 5.4$ PLN, and the new demand $50 + 12\% \cdot 50 = 56$ tons. Before the change, the revenue was $P_1 = C_1 \cdot S_1 = 6 \cdot 50 = 300$ thousand PLN, and the cost $K_1 = 100 + 3S_1 = 100 + 3 \cdot 50 = 250$ thousand PLN. The profit was then equal to $300 - 250 = 50$ thousand PLN. After the change, the revenue reached the value $P_2 = C_2 \cdot S_2 = 5.4 \cdot 56 = 302.4$ thousand PLN, and the cost $K_2 = 100 + 3S_2 = 100 + 3 \cdot 56 = 268$ thousand PLN. Thus the profit is now equal to $302.4 - 268 = 34.4$ thousand PLN.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Example 6
The amount of company's customers in the year t ($t = 0$ for year 2005) is described with sequence (k_t). It is known that on average every second customer attracts a new customer the following year. On the other hand, an average of 20% of the customers gives up after one year and 12% after two years. In 2005, the number of customers amounted to 110 thousand, in 2006 – 121 thousand. Find the formula of (k_t) and the expected number of customers in 2017.

Solution
From the above information it follows that the number of customers in year t equals to the number of customers in year $(t-1)$ increased by 50% of the amount of customers in year $(t-1)$ (new customers), decreased by 20% of the amount of customers in year $(t-1)$ (the customers that gave up after one year) and decreased by 12% of the customers from the year $(t-2)$ (customers that gave up after two years). These conditions can be written as:

$$k_t = k_{t-1} + 0.5k_{t-1} - 0.2k_{t-1} - 0.12k_{t-2}.$$

It leads us to the difference equation:

$$k_t - 1.3k + 0.12k_{t-2} = 0\,.$$

The characteristic equation has the form:

$$x^2 - 1.3x + 0.12 = 0\,.$$

It has two solutions: $x_1 = 0.1$ and $x_2 = 1.2$. It follows that the sequence k_t is described with the formula:

$$k_t = C_1 \cdot 0.1^t + C_2 \cdot 1.2^t\,.$$

By substituting the initial conditions ($k_0 = 110$, $k_1 = 121$), we can find the constants: $C_1 = 10$, $C_2 = 100$. Finally the sequence has the form:

$$k_t = 10 \cdot 0.1^t + 100 \cdot 1.2^t\,.$$

Thus the expected amount of customers in 2017 is equal to:

$$k_{12} = 10 \cdot 0.1^{12} + 100 \cdot 1.2^{12} \approx 891.61 \text{ thousand.}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXERCISES

1. Write the relevant differential equation and find the function having:
   a) constant marginal value b;
   b) constant elasticity b.
2. Find the saturation levels of the functions:
   a) $y = \dfrac{130x}{x+15}$;
   b) $y = \dfrac{220x}{x+7}$;
   c) $y = \dfrac{130(x-7)}{x+15}$;
   d) $y = \dfrac{220(x-11)}{x+7}$.
3. The company has established the following relationship between the demand for sausages, measured in terms of sales S (tons per week), and the average income of consumers D (in PLN for 1 person per month) and the price of sausages C: $S = 30\,D^{0.5}\,C^{-1.5}$. Also the relationship between the weekly production cost K (thousand PLN) and production volume Q (in tons per week) has been estimated: $K = 40 + 10Q$. We assume that the weekly production is adapted to the volume of sales. Current sales $S_1 = 12$ tons, price $C_1 = 15$ PLN/kg.
   a) Calculate the approximate change of demand for sausages, if its price increases by 8%.
   b) Determine how in this case will change weekly revenue, cost and profit.
   c) Specify whether the decision taken is right and why.
4. The company has established the following relationship between the demand for washing powder, measured in terms of sales S (tons per week), and the average income of consumers D (in PLN for 1 person per month) and the price of powder C: $S = 120\,D^{0.6}\,C^{-1.0}$. Also the relationship between the weekly production cost K (thousand PLN) and production volume Q (in tons per week) has been estimated: $K = 100 + 3Q$. Current price is $C_1 = 5$ PLN/kg, current sales $Q_1 = 100$ ton.
   a) Calculate the approximate change of demand for powder, if its price increases by 5%.
   b) Determine how in this case will change weekly revenue, cost and profit.
   c) Specify whether the decision taken is right and why.
5. Demand for good D (in thousand pieces) depends on its price P (PLN) and advertising expenses R (thousand PLN). It is known that 1%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

increase of price is followed by 0.5% decrease of demand, while 1% increase of advertising causes 2% increase of demand. When the price is 25 PLN and the advertising expenses are 10 thousand PLN, the demand equals to 200 thousand pieces. What should be the price if the advertising expenses are 100 thousand PLN and we want the demand to be 400 thousand pieces?

a) Find the demand function.
b) What should be the advertising expenses if the price is 100 PLN and we want the demand to be 400 thousand pieces?

6. The amount of company's customers in the year t ($t = 0$ for year 2010) is described with sequence (a_t). On average every fourth customer attracts a new customer the following year. Find the formula of (a_t) and the expected number of customers in 2017, if:
   a) 5% of the customers give up after one year, 11% after two years, the number of customers in 2010 was 60 thousand, and in 2011 – 56 thousand;
   b) 10% of the customers give up after one year, 5.5% after two years, the number of customers in 2010 was 60 thousand and in 2011 – 13.5 thousand.

SOLUTIONS

1. a) equation: $\frac{dy}{dx} = b$, function: $y = bx + C$ (linear);
   b) equation: $\frac{x}{y} \cdot \frac{dy}{dx} = b$, function: $f(x) = Ce^{bx}$ (power).
2. a) 130; b) 220; c) 130; d) 220.
3. a) the demand will decrease by about 12% (to 10.56 tons);
   b) P_1 = 180,000 PLN, K_1 = 160,000 PLN, Z_1 = 20,000 PLN, P_2 = 171,072 PLN, K_2 = 145,600 PLN, Z_2 = 25,472 PLN;
   c) the decision was right – the profit increased.
4. a) the demand will decrease by about 5% (to 95 tons);
   b) P_1 = 500,000 PLN, K_1 = 400,000 PLN, Z_1 = 100,000 PLN, P_2 = 498,750 PLN, K_2 = 385,000 PLN, Z_2 = 113,750 PLN;
   c) the decision was right – the profit increased.
5. a) $D(P, R) = 200(\mathrm{P}/25)^{-0.5}(R/10)^2$; b) $R = 20$.
6. a) equation: $a_t - 1.2a_{t-1} + 0.11a_{t-2} = 0$,
   sequence: $a_t = 10 \cdot 0.1^t + 50 \cdot 1.1^t$,
   amount of customers in the year 2017: a_7 = 97.44 thousand;
   b) equation: $a_t - 1.15a_{t-1} + 0.055a_{t-2} = 0$,
   sequence: $a_t = 50 \cdot 0.05^t + 10 \cdot 1.1^t$,
   amount of customers in the year 2017: a_7 = 19.5 thousand.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 13
FINANCIAL MATHEMATICS

13.1. COMPOUND INTEREST, STRAMS OF MONEY AND IRR

THEORY IN A NUTSHELL

The value of money changes over time. Whether you will invest your capital in the bank, or leave in your pocket and wait, its value will fall due to inflation: 100 PLN today is worth more than 100 PLN in one year.
The present value of capital will be denoted by *PV*, the future value by *FV*, the number of periods by *n*, and by *i* the interest rate per one period (i.e. the percentage by which the value of capital increases as the period, which can be, for example, interest rate on deposit, but also the rate of return on investment or loan rate). The relationship between the current and future value describes the formula:

$$FV = (1+i)^n PV .$$

It can be transformed to the form:

$$PV = \frac{FV}{(1+i)^n} .$$

Compounding is increasing the principal by the interest (after compounding new interest is calculated based upon the principal and the previously calculated interest). The more frequent is the compounding, the higher is the **effective interest rate**. In the above formulae the number of periods defines also the number of times the compounding took place.
If we sum up several payments, we will obtain the so-called **stream of payments**. The future value of a stream (i.e. the **future value of an annuity** if the payments are equal) is defined by the formula:

$$FV = (1+i)^n PV_1 + (1+i)^{n-1} PV_2 + ... + (1+i)PV_n = \sum_{k=1}^{n} (1+i)^{n-k+1} PV_k .$$

Here we assume that the payments are made at the beginning of each period (i.e. from above), and the value of the stream is calculated in the end of the last period.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

If we want to calculate the present value of future payments made in the end of n periods, i.e. from below (the **present value of an annuity** if the payments are equal), we use the following formula:

$$PV = \frac{FV_1}{(1+i)} + \frac{FV_2}{(1+i)^2} + \ldots + \frac{FV_n}{(1+i)^n} = \sum_{k=1}^{n} \frac{FV_k}{(1+i)^k}.$$

If we want to calculate the present value of infinitely many future payments made in the end of periods (the **present value of a perpetuity** if the payments are equal), we use the following formula:

$$PV = \frac{FV_1}{(1+i)} + \frac{FV_2}{(1+i)^2} + \ldots + \frac{FV_n}{(1+i)^n} + \ldots = \sum_{k=1}^{\infty} \frac{FV_k}{(1+i)^k}.$$

If we know both the future payments and the present value, we can calculate the risk on the investment. The value that we want to find is the **internal rate of return (IRR)**. We find it by solving the following equation with respect to i:

$$PV = \frac{FV_1}{(1+i)} + \frac{FV_2}{(1+i)^2} + \ldots + \frac{FV_n}{(1+i)^n} = \sum_{k=1}^{n} \frac{FV_k}{(1+i)^k}.$$

EXAMPLES

Example 1
We assume that the principal is PV and it increases by i/100% in each period. Find the formula for the value of capital V_n after n periods of compounding.

Solution
It follows that $V_n = (1 + i)V_{n-1}$, or $V_n - (1 + i)V_{n-1} = 0$. The solution of this difference equation is $V_n = C\,(1 + i)^n$. Since $V_0 = PV$, the constant equals to $C = PV$, hence $V_n = PV\,(1 + i)^n$. This way we proved the formula for the future value.

Example 2
We make a deposit of 1,000 PLN. The interest rate is 40% subject to quarterly compounding. What will be value of deposit after 1.5 years?

Solution
The number of quarters in 1.5 years is $n = 6$. The quarterly interest rate is one fourth of the annual one: $i = 0.4/4 = 0.1$. The present value is $PV = 1{,}000$. It follows that the future value of the deposit is:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$FV = (1+0.1)^6 \cdot 1{,}000 = 1{,}771.56 .$$

After 1.5 years we will be able to withdraw 1,771.56 PLN.

Example 3
We want to be able to withdraw 2,000 PLN in two years. The interest rate is 10% subject to semi-annual compounding. How much must we invest?

Solution
The number of semesters in two years is $n = 4$. The semi-annual interest rate is $i = 0.05$, and the future value of deposit is supposed to be $FV = 2{,}000$. It follows that:

$$PV = \frac{2{,}000}{(1+0.05)^4} = 1{,}645.40 .$$

It means that we have to invest 1,645.40 PLN.

Example 4
We invest today 1,000 PLN and 200 PLN less each quarter (payments from above). The interest rate is 20% yearly subject to quarterly compounding. How much will we collect till the end of year?

Solution
The respective stream of payments has been presented in Fig. 13.1.

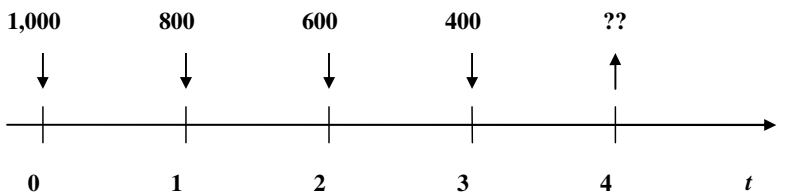


Fig. 13.1. Stream of payments

Quarterly interest rate is $i = 0.05$, the number of quarters is $n = 4$, and the payments in the following quarters are respectively $PV_1 = 1{,}000$, $PV_2 = 800$, $PV_3 = 600$ and $PV_4 = 400$. It follows that the future value of the payments equals to:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$FV = (1+0.05)^4 \cdot 1{,}000 + (1+0.05)^3 \cdot 800 + (1+0.05)^2 \cdot 600 + (1+0.05)^1 \cdot 400 =$$
$$= 3{,}223.11.$$

Thus we will collect this way 3,223.11 PLN till the end of year.

Example 5
The investment will bring 1,000 PLN in one year and 2,000 PLN in two years. What is the maximum reasonable value of investment, if the risk-free interest rate equals to 5%?

Solution
We have to find the present value of the investment. We will use the risk–free rate to estimate, how much we would be able to earn without risk. The respective stream of payments has been presented in Fig. 13.2.

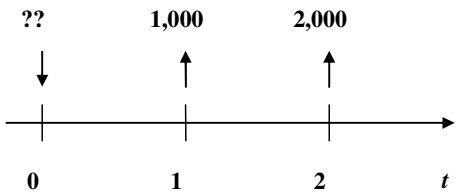


Fig. 13.2. Stream of payments

The interest rate is equal to $i = 0.05$, the number of years is $n = 2$, and the future payments are $FV_1 = 1{,}000$ and $FV_2 = 2{,}000$. By using the appropriate formula we obtain:

$$PV = \frac{1{,}000}{(1+0.05)} + \frac{2{,}000}{(1+0.05)^2} = 2{,}766.44\,.$$

Thus we can earn the given amounts of money without risk, investing today 2,766.44 PLN. It means that 2,766.44 PLN is the highest amount that we should invest.

Example 6
What is the value of deposit that should be made today, if the annual interest rate is 5% and we want to withdraw indefinitely 20,000 PLN in the end of each year?

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution
The example seems similar to the previous one. The main difference is that the payments are supposed to be made indefinitely. The respective stream of payments has been presented in Fig. 13.3.

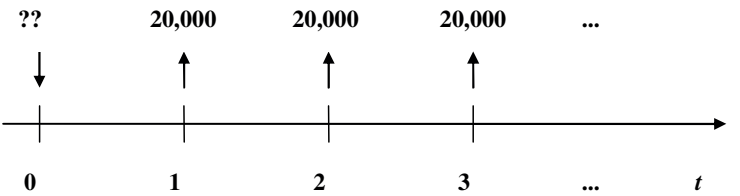


Fig. 13.3. Stream of payments

The interest rate equals to $i = 0.05$, number of payments $n \to \infty$, and the future payments are $FV_1 = 20{,}000$, $FV_2 = 20{,}000$, $FV_3 = 20{,}000$, … In order to find the present value, we use the formula:

$$PV = \frac{20{,}000}{(1+0.05)} + \frac{20{,}000}{(1+0.05)^2} + ... = \sum_{i=1}^{\infty} \frac{20{,}000}{(1+0.05)^i}.$$

The right hand side expression is a geometric series with first term $a_1 = \frac{20{,}000}{1.05}$ and ratio $q = \frac{1}{1.05}$, so its sum equals to:

$$FV = \frac{\frac{20{,}000}{(1+0.05)}}{1-\frac{1}{1.05}} = 400{,}000.$$

It means that to get the desired payments, one has to invest 400,000 PLN.

Example 7
The investment cost is 1,000 PLN today and it brings 500 PLN in one year and 700 PLN in two years. What is the internal rate of revenue?

Solution
Given are the present value $PV = 1{,}000$ and future values $FV_1 = 500$ and $FV_2 = 700$. The equation has the form:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$1{,}000 = \frac{500}{(1+i)} + \frac{700}{(1+i)^2}.$$

Let us substitute $s = 1 + i$ (s is the so-called **growth factor**). We assume that $i \geq -1$ (it is impossible to lose more than 100%), so $s \geq 0$. The equation takes the form:

$$1{,}000 = \frac{500}{s} + \frac{700}{s^2}.$$

By multiplying both sides by s^2 we obtain:

$$10s^2 - 5s - 7 = 0.$$

This equation has two solutions: $s_1 \approx -0.6232$ and $s_2 = 1.1232$. Bearing in mind the assumptions we conclude that $s = 1.1232$, hence $i = 0.1232$. It means that IRR = 12.32%.

EXERCISES

1. Find the future value of the deposit:
   a) 5,000 PLN attracting 6% p.a. interest, subject to annual compounding, in 2 years;
   b) 30,000 PLN attracting 5% p.a. interest, subject to semi-annual compounding, in 3.5 years;
   c) 22,000 PLN attracting 6% p.a. interest, subject to monthly compounding, in 2 years and 2 months.
2. Find the amount of money to be invested in order to be able to withdraw:
   a) 25,000 PLN in 2 years; the deposit attracts 25% p.a. interest, subject to annual compounding;
   b) 17,000 PLN in 3 years; the deposit attracts 10% p.a. interest, subject to semi-annual compounding;
   c) 5,000 PLN in 1 year and 4 months; the deposit attracts 3% p.a. interest, subject to monthly compounding.
3. Find the future value of annuity if the payment rules are as follows:
   a) 100 PLN monthly from above during 12 months, p.a. interest 6% and monthly compounding;
   b) 1,000 PLN quarterly from above during 4.5 years p.a. interest 4% and quarterly compounding;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

c) 500 PLN yearly from above during 20 years, p.a. interest 5% and yearly compounding.

4. Find the present value of the investment, if:
   a) it will bring 10,000 PLN in one year and 20,000 PLN in two years, and the risk-free rate is10%;
   b) it will bring 7,000 PLN in one year and 13,000 PLN in two years, and the risk-free rate is 8%;
   c) it will bring 5,000 PLN in one year, 6,000 PLN in two years and 12,000 PLN in three years, and the risk-free rate is 5%.
5. How much money one needs to collect in order to be able to withdraw indefinitely:
   a) 30,000 PLN in the end of each year, if the interest rate is 7% p.a. (yearly compounding);
   b) 10,000 PLN in the end of each quarter, if the interest rate is 8% p.a. (quarterly compounding);
   c) 1,500 PLN in the end of each month, if the interest rate is 6% p.a. (monthly compounding)?
6. Find the IRR of the investment, which:
   a) costs 10,000 PLN today and will bring 7,000 PLN in one year and 6,600 PLN in two years;
   b) costs 5,000 PLN today and will bring 2,500 in one year and 3,300 PLN in two years;
   c) costs 30,000 PLN today and will bring 13,000 PLN in one year, 12,000 PLN in two years and 11,000 PLN in three years.

SOLUTIONS

1.	a) 5,618 PLN;	b) 35,660.57 PLN;	c) 25,046.11 PLN.
2.	a) 16,000 PLN;	b) 12,685.66 PLN;	c) 4,804.19 PLN.
3.	a) 1,239.72 PLN;	b) 19,810.90 PLN;	c) 17,359.63 PLN.
4.	a) 25,619.83 PLN;	b) 17,626.89 PLN;	c) 20,570.13 PLN.
5.	a) 428,571.43 PLN;	b) 500,000 PLN;	c) 300,000 PLN.
6.	a) 23.46%;	b) 10%;	c) 10%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

13.2. LOAN REPAYMENT SCHEDULE

THEORY IN A NUTSHELL

In order to establish a repayment plan of a loan, you should determine the installments. Once again, we use the equation:

$$PV = \frac{FV_1}{(1+i)} + \frac{FV_2}{(1+i)^2} + ... + \frac{FV_n}{(1+i)^n} = \sum_{k=1}^{n} \frac{FV_k}{(1+i)^k}.$$

This time we solve it with respect to FV_i. We are looking for the installments, for which the loan is fully repaid at the assumed interest rate and repayment period.
There are two types of loans: loans with decreasing installments and loans with fixed installments.
In the case of **decreasing installments**, first we determine the part of principal that is supposed to be paid each period (to do that we divide the total principal by the number of periods and modify the last or the first installment if necessary – rounding errors may occur), then for each period we determine the interest (and the interest is the decreasing part of the payments). The sum of principal and interest is the installment.
In the case of **fixed installments**, first we solve the above mentioned equation, assuming that the installments are equal:

$$PV = \frac{FV}{(1+i)} + \frac{FV}{(1+i)^2} + ... + \frac{FV}{(1+i)^n} = \sum_{k=1}^{n} \frac{FV}{(1+i)^k} = \frac{FV}{i} \cdot \left(1 - \frac{1}{(1+i)^n}\right).$$

The solution is:

$$FV = \frac{PV \cdot i \cdot (1+i)^n}{(1+i)^n - 1}.$$

Then, for each period are calculated: the interest; part of the principal to be paid in this period, determined as the difference between the fixed instalment and the interest; and part of the principal that remains in the end of the period, being equal to the principal in the beginning of the next period. In the last period, the remaining principal is added to the interest. In this way, we determine the value of the last installment. It may be slightly different from the previous ones, because of rounding errors.
Remark. Even if it is not explicitly stated, in the case of loans, compounding takes place each time when one pays an installment.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

EXAMPLES

Example 1

A loan of 100,000 PLN at 20% p.a. has been taken out for 5 years. It is supposed to be repaid in decreasing yearly installments. Present the repayment schedule.

Solution

The principal equals to $PV = 100{,}000$ PLN. It follows that the principal portion in each installment should be equal to 20,000 PLN. The interest rate is $i = 0.2$. The repayment schedule has been presented in Table 13.1.

Table 13.1. Repayment schedule (in PLN)

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	100,000.00	20,000.00	20,000.00	20,000.00	40,000.00	80,000.00
2	80,000.00	16,000.00	16,000.00	20,000.00	36,000.00	60,000.00
3	60,000.00	12,000.00	12,000.00	20,000.00	32,000.00	40,000.00
4	40,000.00	8,000.00	8,000.00	20,000.00	28,000.00	20,000.00
5	20,000.00	4,000.00	4,000.00	20,000.00	24,000.00	0.00

Let us look at the **first row**. The beginning principal is 100,000 PLN. The interest is equal to $0.2 \cdot 100{,}000$, i.e. 20,000 PLN. The interest is supposed to be fully paid, so the interest portion of the payment is 20,000 PLN. The principal portion also equals to 20,000 PLN, as we already know. Thus the total payment is 20,000 + 20,000 = 40,000 PLN (this is the amount to be paid to the bank). The ending principal is the beginning principal decreased by the principal portion (not by the total payment!). Thus it is equal to 100,000 – 20,000 = 80,000 PLN. It is the beginning principal for the next period.

In the **second row**, the beginning principal is 80,000 PLN. The interest is equal to $0.2 \cdot 80{,}000$, i.e. 16,000 PLN. The interest is supposed to be fully paid, so the interest portion of the payment is 16,000 PLN. The principal portion equals to 20,000 PLN, so the total payment is 16,000 + 20,000 = 36,000 PLN. The ending principal is the beginning principal decreased by the principal portion, i.e. 80,000 – 20,000 = 60,000 PLN. It is the beginning principal for the next period. We continue this way until the ending principal becomes 0 PLN.

Example 2

A loan of 100,000 PLN at 20% p.a. has been taken out for 5 years. It is supposed to be repaid in fixed yearly installments. Present the repayment schedule.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Solution

The principal is equal to $PV = 100{,}000$ PLN, but this time we cannot find the principal portions of the installments based only on this information. The interest rate is $i = 0.2$. Since the loan is supposed to be paid in equal installments, we derive their value from the formula:

$$100{,}000 = \frac{R}{(1+0.2)} + \frac{R}{(1+0.2)^2} + \frac{R}{(1+0.2)^3} + \frac{R}{(1+0.2)^4} + \frac{R}{(1+0.2)^5}.$$

The expression on the right is the sum of geometric sequence having $n = 5$ terms, first term $a_1 = \frac{R}{1.2}$ and common ratio $q = \frac{1}{1.2}$. Thus it can be reduced to the form:

$$100{,}000 = \frac{R}{1.2} \cdot \frac{1 - \frac{1}{1.2^5}}{1 - \frac{1}{1.2}}.$$

It follows that the installments are equal to:

$$R = \frac{100{,}000 \cdot 1.2^5 \cdot 0.2}{1.2^5 - 1} = 33{,}437.97.$$

The repayment schedule has been presented in Table 13.2.

Table 13.2. Repayment schedule (in PLN)

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	100,000.00	20,000.00	20,000.00	13,437.97	33,437.97	86,562.03
2	86,562.03	17,312.41	17,312.41	16,125.56	33,437.97	70,436.47
3	70,436.47	14,087.29	14,087.29	19,350.68	33,437.97	51,085.79
4	51,085.79	10,217.16	10,217.16	23,220.81	33,437.97	27,864.98
5	27,864.98	5,573.00	5,573.00	27,864.98	33,437.98	0.00

The computations in the first four rows are conducted in the same way (note that their order is different than in the case of decreasing installments). Let us look at the **first row**. The beginning principal is equal to 100,000 PLN, so the interest is $0.2 \cdot 100{,}000 = 20{,}000$ PLN, and as it has to be fully paid, the interest portion of the installment is 20,000 PLN (to this moment the computations are performed in the same order as in the case of decreasing installments). We do not know the principal portion, but we know that the total payment is 33,437.97 PLN. Since

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

the interest portion is 20,000 PLN, finally we can allocate 33,437.97 – 20,000 = 13,437.97 PLN to the principal (it is the principal portion of the loan). It means that the ending principal is equal to 100,000 – 13,437.97 = 86,562.03 PLN and it becomes the beginning principal for the next period. In the **second row**, the beginning principal equals to 86,562.03 PLN, hence the interest is $0.2 \cdot 86{,}562.03 = 17{,}312.41$ PLN, and since it has to be fully paid, the interest portion equals to 17,312.41 PLN. The total payment equals to 33,437.97 PLN. Since the interest portion is 17,312.41 PLN, finally we can allocate 33,437.97 – 17,312.41 = 16,125.56 PLN to the principal. It means that the ending principal is 86,562.03 – 16,125.56 = 70,436.47 PLN, which is also the beginning principal for the next period. The computations are the same for all the periods except the last one. In the **last period**, we do not assume that the total payments is given – we calculate it as in the case of decreasing installments. The beginning principal equals to 27,864.98 PLN, hence the interest is $0.2 \cdot 27{,}864.98 = 5{,}573$ PLN, and since it has to be fully paid, the interest portion equals to 5,573 PLN. We are analyzing the last period, so the principal portion must be equal to all the remaining principal, so it equals to 27,864.98 PLN. Hence the total payment in the last period equals to 5,573 + 27,864.98 = 33,437.98 PLN. Note that the last installment differs by 0.01 PLN from the other ones – this is due to the rounding errors that accumulate over all periods. Do not worry if in your calculations, this difference will come out a bit more; it may be of the order to few cents, if you are rounding the amounts to two decimal places.

EXERCISES

1. Present the loan repayment schedule, if the installments are decreasing and the remaining conditions are:
    a) principal 100,000 PLN, 5 years, yearly payments, interest 5% p.a.;
    b) principal 200,000 PLN, 1 year, quarterly payments, interest 4% p.a.;
    c) principal 10,000 PLN, 1 year, monthly payments, interest 2.4% p.a.
2. Present the loan repayment schedule, if the installments are fixed and the remaining conditions are:
    a) principal 100,000 PLN, 5 years, yearly payments, interest 5% p.a.;
    b) principal 200,000 PLN, 1 year, quarterly payments, interest 4% p.a.;
    c) principal 10,000 PLN, 1 year, monthly payments, interest 2.4% p.a.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

SOLUTIONS

1. Repayment schedules:

a)

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	100,000.00	5,000.00	5,000.00	20,000.00	25,000.00	80,000.00
2	80,000.00	4,000.00	4,000.00	20,000.00	24,000.00	60,000.00
3	60,000.00	3,000.00	3,000.00	20,000.00	23,000.00	40,000.00
4	40,000.00	2,000.00	2,000.00	20,000.00	22,000.00	20,000.00
5	20,000.00	1,000.00	1,000.00	20,000.00	21,000.00	0.00

b)

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	200,000.00	2,000.00	2,000.00	50,000.00	52,000.00	150,000.00
2	150,000.00	1,500.00	1,500.00	50,000.00	51,500.00	100,000.00
3	100,000.00	1,000.00	1,000.00	50,000.00	51,000.00	50,000.00
4	50,000.00	500.00	500.00	50,000.00	50,500.00	0.00

c)

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	10,000.00	20.00	20.00	833.33	853.33	9166.67
2	9,166.67	18.33	18.33	833.33	851.66	8333.34
3	8,333.34	16.67	16.67	833.33	850.00	7500.01
4	7,500.01	15.00	15.00	833.33	848.33	6666.68
5	6,666.68	13.33	13.33	833.33	846.66	5833.35
6	5,833.35	11.67	11.67	833.33	845.00	5000.02
7	5,000.02	10.00	10.00	833.33	843.33	4166.69
8	4,166.69	8.33	8.33	833.33	841.66	3333.36
9	3,333.36	6.67	6.67	833.33	840.00	2500.03
10	2,500.03	5.00	5.00	833.33	838.33	1666.70
11	1,666.70	3.33	3.33	833.33	836.66	833.37
12	833.37	1.67	1.67	833.37	835.04	0.00

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. Equal installments and the repayment schedules:

a) installment 23,097.48 PLN;

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	100,000.00	5,000.00	5,000.00	18,097.48	23,097.48	81,902.52
2	81,902.52	4,095.13	4,095.13	19,002.35	23,097.48	62,900.17
3	62,900.17	3,145.01	3,145.01	19,952.47	23,097.48	42,947.70
4	42,947.70	2,147.39	2,147.39	20,950.09	23,097.48	21,997.61
5	21,997.61	1,099.88	1,099.88	21,997.61	23,097.49	0.00

b) installment 51,256.22 PLN;

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	200,000.00	2,000.00	2,000.00	49,256.22	51,256.22	150,743.78
2	150,743.78	1,507.44	1,507.44	49,748.78	51,256.22	100,995.00
3	100,995.00	1,009.95	1,009.95	50,246.27	51,256.22	50,748.73
4	50,748.73	507.49	507.49	50,748.73	51,256.22	0.00

c) installment 844.21 PLN.

Period	Beginning principal	Interest accrued	Interest portion	Principal portion	Total payment	Ending principal
1	10,000.00	20.00	20.00	824.21	844.21	9,175.79
2	9,175.79	18.35	18.35	825.86	844.21	8,349.93
3	8,349.93	16.70	16.70	827.51	844.21	7,522.42
4	7,522.42	15.04	15.04	829.17	844.21	6,693.25
5	6,693.25	13.39	13.39	830.82	844.21	5,862.43
6	5,862.43	11.72	11.72	832.49	844.21	5,029.94
7	5,029.94	10.06	10.06	834.15	844.21	4,195.79
8	4,195.79	8.39	8.39	835.82	844.21	3,359.97
9	3,359.97	6.72	6.72	837.49	844.21	2,522.48
10	2,522.48	5.04	5.04	839.17	844.21	1,683.31
11	1,683.31	3.37	3.37	840.84	844.21	842.47
12	842.47	1.68	1.68	842.47	844.15	0.00

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

CHAPTER 14
PROBABILITY

14.1. BASIC CONCEPTS

THEORY IN A NUTSHELL

Definition 14.1. Event
Event is the outcome of an experiment, about the result of which we cannot be sure.

The examples of events are: obtaining at most 2 by a throw of one dice, getting heads in one flip of a coin or drawing one of the first five out of thirty questions during the diploma exam.

Definition 14.2. Simple event
Simple event is an outcome which cannot be broken down.

The examples of simple events are: obtaining exactly two by the throw of one dice, getting heads in one flip of a coin or drawing the third question during the diploma exam.

Definition 14.3. Sample space
Sample space Ω is the set of all the simple events.

Each event consists of simple events and thus is a subset of the sample space. As the events are sets, we can perform on them all the set operations (as union or intersection, see Chapter 2). Also, all the set algebra laws are applicable.

Definition 14.4. Probability
Probability $P(A)$ is the measure of the likeliness that an event A will occur.

For example the probability of obtaining 2 in a throw of one dice is 1/6, because it is one of the six equally likely events. Similarly getting heads in one flip of a coin is an event occurring with probability 1/2, and drawing the third question out of thirty during the diploma exam occurs with probability 1/30.
The basic properties of probability have been presented below. Here A and B denote arbitrary events, Ω – sample space (or **certain event**), $\varnothing$ – event that will never occur (**impossible event**):

1. $P(A) \geq 0$, $P(A) \leq 1$.
2. $P(\varnothing) = 0$, $P(\Omega) = 1$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

3. Probability of the complement of an event: $P(A') = 1 - P(A)$.
4. Probability of the union of two events (Inclusion–Exclusion Principle): $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

The last property can be generalized for any number of events, just as in the case of the Inclusion–Exclusion Principle for sets (see Chapter 2).

Definition 14.5. Conditional probability
The probability of an event A given that another event B has occurred equals to:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}.$$

The conditional probability $P(A / B)$ (we read it: probability of A given B) that event A will occur given that event B has occurred (do not confuse it with the probability of the difference of two events $P(A \setminus B)$) is the probability of A, when we already know, that B has occurred. An example is the probability of obtaining 2 on a dice, given that the obtained number is less than or equal to 4 – in such a situation only four possibilities have left, so the probability is 1/4.
It is often the case that one knows the conditional probabilities of an event A given $B_1, B_2, \ldots, B_n$, but the probability of A is unknown. In some cases (when $B_1, B_2, \ldots, B_n$ are mutually disjoint and their sum is the sample space) we can use the Law of Total Probability.

Theorem 14.1. Law of Total Probability
Assume that the events $B_1, B_2, \ldots, B_n$ satisfy the conditions:
1. $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$.
2. $B_i \cap B_j = \varnothing$ for $i \neq j$.

Then the following is true:

$$P(A) = \sum_{k=1}^{n} P(A/B_k)P(B_k) =$$
$$= P(A/B_1)P(B_1) + P(A/B_2)P(B_2) + \ldots + P(A/B_n)P(B_n).$$

Sometime one knows the conditional probabilities of an event A given $B_1, B_2, \ldots, B_n$, but they want to know the probability of one of the events B_j given A. Then the **Bayes Law** is helpful.

Theorem 14.2. Bayes Law
Assume that the events $B_1, B_2, \ldots, B_n$ satisfy the conditions:
1. $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$.
2. $B_i \cap B_j = \varnothing$ for $i \neq j$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Then the following is true:

$$P(B_j / A) = \frac{P(A/B_j)P(B_j)}{P(A)} = \frac{P(A/B_j)P(B_j)}{\sum_{k=1}^{n} P(A/B_k)P(B_k)}.$$

EXAMPLES

Example 1

Transportation company has three trucks: S_1, S_2 and S_3. Truck S_1 makes 50% of supplies, truck S_2 – 30%, and truck S_3 – 20%. Truck S_1 fails on average once every 100 courses, S_2 two times per 100 courses, and S_3 three times per 100 courses.

1. What is the probability that a randomly selected course will be trouble-free?
2. One of the trucks failed. What are the chances that it is S_2?

Solution

We need to determine which of the events are the cause and which effect. This is done on the basis of information on conditional probabilities. As you can see, there are certain probabilities of failure, depending on the vehicles. This means that the cause is the departure of the vehicle on the road, while the effect is a failure (this can also be deduced directly: first vehicle going on a tour, then a failure occurs or not). So we denote by A the event consisting in the fact that the failure occurred, and by B_1, B_2 and B_3 – events based on the fact that the course was held by the truck S_1, S_2 and S_3, respectively. From the conditions it follows that:

- $P(B_1) = 0.5$; truck S_1 holds 50% of courses, so the chances that a randomly chosen course is held by this truck are 50%; analogously $P(B_2) = 0.3$ and $P(B_2) = 0.2$;
- $P(A/B_1) = 0.01$; is the probability that a vehicle has failed, if we know that it was the truck S_1; this vehicle fails on average once every 100 courses, so the probability is 0.01; analogously $P(A/B_2) = 0.02$ and $P(A/B_3) = 0.03$.

Now we have to check whether the assumptions of the Law of Total Probability and Bayes Law are satisfied. Since the company has only three vehicles, the randomly chosen course must have been held by one of the trucks S_1 or S_2 or S_3. Hence $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$. Moreover, a course cannot be held simultaneously by two cars, thus $B_i \cap B_j = \varnothing$ for $i \neq j$. Now we can determine the probabilities:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Ad 1:

$$P(A) = P(A/B_1)P(B_1) + P(A/B_2)P(B_2) + P(A/B_3)P(B_3) =$$
$$= 0.01 \cdot 0.5 + 0.02 \cdot 0.3 + 0.02 \cdot 0.2 = 0.017.$$

Thus the probability of a failure equals to 0.017. It follows that the probability that the course will be trouble-free (i.e. the probability of the complement of A) is $P(A') = 1 - 0.017 = 0.983$.

Ad 2:

$$P(B_2/A) = \frac{P(A/B_2)P(B_2)}{P(A)} = \frac{0.02 \cdot 0.3}{0.017} = 6/17\,.$$

As you can see, the probability that the car which failed is S_2, equals to 6/17.

Example 2
Player in a TV game show has a choice of three gates. There is a prize in one of them. When the player chooses a gate (regardless of whether containing the prize, or empty), the presenter exposes one of the other two gates, always empty, and then gives the player a choice – to stay with the first gate, or to choose another. What should the player do?

Solution
The player does not know where is the prize. Thus they should calculate the probability of winning for each choice and choose the gate with the highest probability.
This time it is a little harder to determine what is cause and what is effect. Let us establish first what is the event.
Surely the presence of the prize in one of the gates is an event. The choice of a gate is not (player chooses a gate intentionally, and twice). However, from the point of view of the player, the choice of a gate by the presenter is an event (of course, the presenter also chooses a gate intentionally, but the player knows neither selection rules, nor the contents of the gates, so it can be assumed that the choice is random).
What is the cause and what is the effect? The presenter chooses a gate based on information about where is the prize (cannot open the gate with the prize), so the causes are the events consisting of the presence of the prize in the gates, and the result is the choice made by the presenter.
Let us number the gates. The gate chosen by the player will be denoted by 1, and the gate opened by the presenter by 3. This will not reduce the generality, since we can repeat the reasoning for each numbering. The player must then choose

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

between the gate number 1 (stay with the original choice) and gate number 2 (change the decision). Let us denote the events:

B_1 – the prize is in the gate number 1;
B_2 – the prize is in the gate number 2;
B_3 – the prize is in the gate number 3 (yes! We have to take into account this event; we already know that there is no prize in this gate – it was chosen by the presenter – but the player does not know this at the beginning!);
A – the presenter choose the gate number 3.

Since the gate, where the prize would be placed, is drown, we have:

$P(B_1) = P(B_2) = P(B_3) = 1/3$.

Let us determine the conditional probabilities. If the prize is in the gate number 1 (i.e. the one chosen by the player), the presenter can open each of the two remaining gates (they are both empty). Thus the probability that they open the gate number 3 is $P(A/B_1) = 1/2$. If the prize is in the gate number 2, the presenter has no choice – they have to open the gate number 3, so the probability is $P(A/B_2) = 1$. If the prize is in the gate number 3, the presenter cannot open it, so the probability is $P(A/B_3) = 0$.
We know that the presenter opened the gate number 3. The probabilities that we are looking for are $P(A/B_1)$ (i.e., the probability that the prize is in the gate chosen by the player) and $P(A/B_2)$ (i.e., the probability, that the prize is in the other gate):

$$P(A) = P(A/B_1)P(B_1) + P(A/B_2)P(B_2) + P(A/B_3)P(B_3) = \\ = 1/2 \cdot 1/3 + 1 \cdot 1/3 + 0 \cdot 1/3 = 1/2,$$

$$P(B_1/A) = \frac{P(A/B_1)P(B_1)}{P(A)} = \frac{1/2 \cdot 1/3}{1/2} = 1/3,$$

$$P(B_2/A) = \frac{P(A/B_2)P(B_2)}{P(A)} = \frac{1 \cdot 1/3}{1/2} = 2/3.$$

As you can see, the winning chances are two times higher after changing the choice. Thus the player should change the gate.

EXERCISES

1. The eggs in a supermarket come from three farms X, Y, Z. We know that 20% come from X, 50% from Y and 30% from Z. It is also known that 2%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

of the eggs from X are infected with salmonella, while this percentage for farms Y and Z is 1% and 3% respectively. What is the probability that:

a) an egg randomly chosen in the supermarket is not infected;
b) an infected egg comes from the farm X.

2. In the IT company X, two developers P_1 and P_2 work together to write an application. The efficiency of P_1 is 600 lines of code per day, the efficiency of P_2 is 300 lines per day. The probability that P_1 makes a mistake in chosen line is 0.05, the same probability for P_2 is 0.15.
   a) What is the probability that there is a mistake in randomly chosen line?
   b) There is a mistake in randomly chosen line. What is the probability that it was made by P_1?
3. In the factory there are four machines of type A, four machines of type B and two of type C. They produce the product with same efficiency. Among the products produced by A there are 10% of faults, while for the machines B and C these ratios are 5% and 15%, respectively.
   a) One product has been chosen at random. What is the probability that it is a fault?
   b) One product has been chosen at random. It is a fault. What is the probability that it was produced by the machine B?
4. We have three coins: one is a usual coin with heads and tails, two others have defects – one has only heads, one has only tails.
   a) We flip randomly chosen coin. What is the probability of heads?
   b) We flip randomly chosen coin, without checking which coin it is. It lands tails. What is the probability that this coin has only tails?
5. Approximately 12% of companies fail after some time. Among the failing companies 80% have very low level of a financial indicator. Among the ones which do not fail, 30% have very low level of this indicator. The bank offers a loan only to the companies that have high level of the indicator.
   a) What is the probability that randomly chosen company will get the loan?
   b) What is the probability of bankruptcy of a company that gets the loan?

SOLUTIONS

1.	a) 0.982;	b) 2/9.
2.	a) 1/12;	b) 0.4.
3.	a) 0.09;	b) 2/9.
4.	a) 0.5;	b) 2/3.
5.	a) 0.64;	b) 3/80.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

14.2. RANDOM VARIABLE

THEORY IN A NUTSHELL

Definition 14.6. Random variable
Random variable is a function defined on the sample space.

In other words, the random variable is a function that assigns to each elementary event a numerical value. For example, every possible result of a dice roll, we can assign a numerical value equal to the number thrown. If we look at this definition, on the other hand, we can assign probabilities to the respective values of the random variable.
We denote the random variables by uppercase letters (for example X), and their values by lowercase letters (for example x, x_1). The probability that variable X equals to a, we denote by $P(X = a)$. Analogously we denote the probabilities that X is a member of given interval or satisfies given inequalities. The set of all the pairs $(x, P(X = x))$ is called the **distribution** of X.
We say that random variable has **discrete distribution**, when we know the probabilities of all its values. Sample discrete distribution is the distribution of the number on a dice: we can attach a probability to each of those numbers. We describe the discrete distributions in one of the following ways:

1. We can list all the values of the random variable end corresponding probabilities. In the case of throwing a dice, such a distribution would take the form $\{(1, 1/6), (2, 1/6), (3, 1/6), (4, 1/6), (5, 1/6), (6, 1/6)\}$.
2. We can use a formula. In the case of throwing a dice, we can use the formula $P(X = k) = 1/6$, for $k = 1, 2, \ldots, 6$.

The sum of the probabilities corresponding with all the values of a random variable **is always equal to 1**.
The distribution with all the probabilities equal is called **uniform distribution**. In addition, we will be interested also in two other types of discrete distributions. First of them is the **binomial distribution**. The experiment consists of repeated conduct of the same trial, in which the probability of occurrence of a specific event (the probability of success) is fixed (we denote it by p). The number of successes in n trials has the binomial distribution with parameters n and p, i.e., the probability of achieving the value k by this variable is given by:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, 1, 2, \ldots, n.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Let us recall that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the **binomial coefficient**, and $n!$ denotes the **factorial**, i.e. the product of all the integers from 1 to n (moreover we have $0! = 1$). The fact that a random variable X has binomial distribution (or is binomially distributed) with parameters n and p is denoted by $\boldsymbol{X \sim B(n, p)}$.
Another important distribution is the **Poisson distribution**. We are dealing with similar situation as above, but the number of trials is huge and we only know the average number of successes λ (lambda). The probability of achieving the value k by the random variable equal to the number of successes is given by:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ for } k = 0, 1, 2, \ldots.$$

The fact that a random variable X has Poisson distribution with parameter λ, is denoted by $\boldsymbol{X \sim Po(\lambda)}$.
The distribution $Po(\lambda)$ is a good approximation of $B(n, p)$ for large n (in practice $n \geq 30$) and very small p (in practice $p < 0.1$ simultaneously with $np < 20$).
Random variable can also have a **continuous distribution**. In such a situation the probability that it achieves any specific value is always equal to 0. However, we can define the probability that it falls into some specific interval. In this case the distribution is defined using the **density function** $f(x)$, and the probability that the variable falls into the interval (a, b), is equal to $\int_a^b f(x)\,dx$ (a may be replaced with $-\infty$, and b with ∞, if necessary).
The continuous variant of the **uniform distribution** with parameters x_1 and x_2 is defined by the density function:

$$f(x) = \frac{1}{x_2 - x_1}, \; x \in [x_1, x_2].$$

Outside the given interval the density function equals to 0.
The fact that X has uniform distribution with parameters x_1 and x_2 is denoted by $\boldsymbol{X \sim U(x_1, x_2)}$.
Random variable, which is **exponentially distributed** with parameter λ, has the density function:

$$f(x) = \lambda e^{-\lambda x}, \; x \geq 0.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

The fact that X has exponential distribution with parameter λ is denoted by $\boldsymbol{X \sim Exp(\lambda)}$.

Very important in economical (and not only) applications is the **normal distribution** (called also **Gaussian distribution**). The density function of normal distribution with parameters μ (mu) and σ (sigma) is as follows:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

The fact that X has normal distribution with parameters μ and σ is denoted by $\boldsymbol{X \sim N(\mu, \sigma)}$.

Note that in the case of $N(\mu, \sigma)$ we use special symbol to denote the density function: φ (phi).

If a random variable X has normal distribution with parameters μ and σ, then the random variable $X_{st} = (X - \mu)/\sigma$ has normal distribution with parameters 0 and 1, i.e. the **standard normal distribution**. The variable X_{st} is the **standardized** X, and the respective operation is called **standardization**. It is very important, because the normal distribution is so frequently used that it has been tabelarized[28]. In the tables you will find, however, only the probabilities corresponding with the distribution $N(0, 1)$.

The normal distribution $N(np, \sqrt{np(1-p)})$ can be used to approximate the binomial distribution $B(n, p)$ for large n (when $n \geq 30$) and relatively large p (when $p \geq 0.1$ or $np \geq 20$). To be more specific, in order to find the probability $P(X \in [a, b])$, it is necessary to read from the tables of normal distribution the probability $P(X \in [a - 0.5, b + 0.5])$, i.e. to calculate the difference $P(X < b + 0.5) - P(X < a - 0.5)$. In order to find the probability $P(X = a)$, you need to find the probability $P(X \in [a - 0.5, a + 0.5])$, i.e. to calculate the difference $P(X < a + 0.5) - P(X < a - 0.5)$.

Based on knowledge of the distribution (probability formula or density function) we ususally cannot say too much about the random variable. Therefore, some **distribution characteristics** have been introduced. Four of them have been listed below.

1. **Cumulative distribution function (CDF)**, denoted by $F(x)$[29] measures the cumulated probability. Its value is equal to the probability that the random variable will be less than given value x[30].

[28] The table together with short guide can be found in Section 14.3.

[29] In the case of normal distribution we use the letter Φ (phi).

[30] Sometimes it is defined as the probability that the random variable will be less than or equal to given value x. It has influence on the formula or value of CDF in the case of continuous variable and in the case of discrete variables changes the range of summation

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

2. **Expected value** (or **expectation**) is the average value of the variable.
3. **Variance** measures the spread of the values of random variables around its expectation, however it has no interpretation.
4. **Standard deviation** is the mean deviation of the variable from its expectation.

The formulae of the above characteristics have been presented in Table 14.1.

Table 14.1. Distribution characteristics

Characteristic	Discrete random variable*	Continuous random variable
CDF	$F(x)=\sum_{x_i<x} P(x_i)$	$F(x)=\int_{-\infty}^{x} f(t)dt$
Expectation	$E(X)=\sum_{x_i} x_i P(x_i)$	$E(X)=\int_{-\infty}^{\infty} xf(x)dx$
Variance	$V(X)=$ $=\sum_{x_i} x_i^2 P(x_i)-\left[E(X)\right]^2$	$V(X)=$ $=\int_{-\infty}^{\infty} x^2 f(x)dx-\left[E(X)\right]^2$
Standard deviation	$S(X)=\sqrt{V(X)}$	$S(X)=\sqrt{V(X)}$

* The summation range: "$x_i < x$" means that we sum the terms for all the values x_i of X less than x, and the range "x_i" means that we sum the terms for all the values x_i of X.

Table 14.2. Characteristics of important distributions*

Characte-ristic	$X \sim B(n, p)$	$X \sim Po(\lambda)$	$X \sim U(x_1, x_2)$	$X \sim Exp(\lambda)$	$X \sim N(\mu, \sigma)$
$F(x)$	——	——	$\frac{x-x_1}{x_2-x_1}$	$1-e^{-\lambda x}$	——
$E(X)$	np	λ	$\frac{x_1+x_2}{2}$	$1/\lambda$	μ
$V(X)$	$np(1-p)$	λ	$\frac{(x_2-x_1)^2}{12}$	$1/\lambda^2$	σ^2
$S(X)$	$\sqrt{np(1-p)}$	$\sqrt{\lambda}$	$\frac{x_2-x_1}{2\sqrt{3}}$	$1/\lambda$	σ

* The formulae have been skipped if they have no simple form. The CDF given only for the interval where $f(x) > 0$.

---

from $x_i < x$ to $x_i \le x$ and the value of CDF changes only in the points which can be achieved by the variable.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In the case of the most popular distributions mentioned above, it makes no sense to derive their characteristics again and again. For that reason they have been presented in Table 14.2.

EXAMPLES

Example 1

Discrete random variable X has distribution: $\{(1, 0.2), (3, 0.7), (5, p)\}$. Determine p, and then find the CDF, expectation and standard deviation of X.

Solution

Since the variable must take one of the three values (1, 3, 5), the sum of corresponding probabilities must be 1: $0.2 + 0.7 + p = 1$. Hence $p = 0.1$. We can now proceed to determine the characteristics of the distribution.

Cumulative distribution function is defined on the set of real numbers. Its value at any point x is equal to the probability that the random variable will be less than x. For 1 and all the smaller numbers this probability equals to 0 (the variable does not take any values less than 1, so also less than any number less than 1), thus $F(x) = 0$, when $x \in (-\infty, 1]$. For the numbers from the interval (1, 3] this probability is 0.2, because the random variable can take only one value less than any of them, i.e. 1, so $F(x) = 0.2$, when $x \in (1, 3]$. For the numbers from the interval (3, 5] this probability equals to $0.2 + 0.7 = 0.9$, because the random variable can take only two values less than them, i.e. 1 or 3, thus $F(x) = 0.9$, when $x \in (3, 5]$. If we choose any number from the interval $(5, \infty)$, the random variable will be surely less than that (it takes value 1, 3 or 5), so the probability will be 1. Finally, the CDF is described by the formula:

$$F(x) = \begin{cases} 0, & \text{if } x \in (-\infty, 1], \\ 0.2, & \text{if } x \in (1, 3], \\ 0.9, & \text{if } x \in (3, 5], \\ 1, & \text{if } x \in (5, \infty). \end{cases}$$

Note that all the intervals are right-closed[31], and their ends are exactly the values taken by the random variable.

Let us proceed to the expectation. According to the formula we have to add the products of the values of the variable and corresponding probabilities:

$$E(X) = \sum_{x_i} x_i P(x_i) = 1 \cdot 0.2 + 3 \cdot 0.7 + 5 \cdot 0.1 = 2.8 .$$

[31] If we used the alternative definition of CDF, the intervals would be left-closed.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

It means that the variable has average value 2.8.
Let us proceed to the standard deviation. In order to do that, first we have to find the variance using the appropriate formula. Its first component is the sum of products of squares of the values taken by the random variable and corresponding probabilities. The second component is the square of expected value:

$$V(X) = \sum_{x_i} x_i^2 P(x_i) - [E(X)]^2 = 1^2 \cdot 0.2 + 3^2 \cdot 0.7 + 5^2 \cdot 0.1 - 2.8^2 = 1.16 .$$

The standard deviation is the square root of the variance:

$$S(X) = \sqrt{V(X)} = \sqrt{1.16} \approx 1.077 .$$

Its value means that the random variable X differs on average by 1.077 from its mean value (i.e., 2.8).

Example 2
Find the CDF, expected value and standard deviation of the variable $X \sim U(1, 6)$.

Solution
Given variable has uniform distribution with parameters 1 and 6, so it is described by the density function:

$$f(x) = \begin{cases} 1/5, \text{ if } x \in [1, 6], \\ 0, \quad \text{if } x \notin [1, 6]. \end{cases}$$

In order to find $F(x)$, we have to find the value of the following integral for every value of x:

$$F(x) = \int_{-\infty}^{x} f(t)dt .$$

We will do it interval-wise as the formula of the density function is different in different intervals. On the interval $(-\infty, 1]$[32] it takes the value 0, so for $x \in (-\infty, 1]$ we have:

[32] We can use closed intervals. Although the density function changes its value at $x = 1$, the value of function at one point does not change the value of the definite integral.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0.$$

Since for each $x \in (-\infty, 1]$ the CDF equals to 0, in particular we have $F(1) = 0$. We will use this fact in the remainder. Let us proceed for $x \in (1, 6]$. This time we have to split the domain of integration – the formula of the density function changes, when passing 1:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{1} f(t)dt + \int_{1}^{x} f(t)dt = \ldots$$

Note that the first integral is by definition equal to

$$\int_{-\infty}^{1} f(t)\,dt = F(1) = 0$$

(it was the reason for computing $F(1)$). Finally we have:

$$\ldots = F(1) + \int_{1}^{x} 1/5\,dt = 0 + [1/5t]_1^x = 1/5x - 1/5.$$

Similarly as above, we compute the value of CDF at the end of the interval: $F(6) = 1/5 \cdot 6 - 1/5 = 1$. The last interval left: $x \in (6, \infty)$. Using similar methods as above, we obtain:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{6} f(t)\,dt + \int_{6}^{x} f(t)\,dt = F(6) + \int_{6}^{x} 0\,dt = 1 + 0 = 1.$$

Note that also this time we subdivided the domain into two parts, although the density function is defined separately on three intervals. It follows from the fact that we found the value of $F(6)$ – thanks to that we already know the sum of two integrals $\int_{-\infty}^{1} f(t)dt + \int_{1}^{6} f(t)dt$. Finally, the formula describing CDF is:

$$F(x) = \begin{cases} 0, & \text{if } x \in (-\infty, 1], \\ 1/5x - 1/5, & \text{if } x \in (1, 6], \\ 1, & \text{if } x \in (6, \infty). \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In order to find the expectation we have to compute the integral:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx .$$

Since the density function is defined independently on three intervals, we subdivide the domain:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{1} xf(x)dx + \int_{1}^{6} xf(x)dx + \int_{6}^{\infty} xf(x)dx =$$
$$= \int_{-\infty}^{1} x \cdot 0dx + \int_{1}^{6} x \cdot 1/5dx + \int_{6}^{\infty} x \cdot 0dx = 0 + [1/10x^2]_1^6 + 0 = 3.5.$$

In similar way we find the variance:

$$V(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - [E(X)]^2 = \int_{-\infty}^{1} x^2 \cdot 0dx + \int_{1}^{6} x^2 \cdot 1/5dx + \int_{6}^{\infty} x^2 \cdot 0dx - 3.5^2 =$$
$$= 0 + [1/15x^3]_1^6 + 0 - 3.5^2 = 25/12.$$

The standard deviation is thus equal to

$$S(X) = \sqrt{V(X)} = \sqrt{25/12} = 5\sqrt{3}/6 .$$

Example 3
Find the CDF of the variable $X \sim Exp(2)$.

Solution
The density function is this time:

$$f(x) = \begin{cases} 2e^{-2x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

We will compute the CDF interval-wise, as in the previous example. For $x \leq 0$ we have:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0 .$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Hence $F(0) = 0$. For $x > 0$ the CDF is in turn equal to:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt = F(0) + \int_{0}^{x} 2e^{-2t}dt = 0 + [-e^{-2t}]_0^x = 1 - e^{-2x}.$$

Finally:

$$F(x) = \begin{cases} 0, & \text{if } x \le 0, \\ 1 - e^{-2x}, & \text{if } x > 0. \end{cases}$$

Example 4
Find the CDF, expectation and standard deviation of the random variable defined by the density function:

$$f(x) = \begin{cases} \dfrac{2}{x^3}, & x \ge 1, \\ 0, & x < 1. \end{cases}$$

Solution
We compute the CDF interval-wise. For $x \le 1$ we have:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0.$$

Hence $F(1) = 0$. For $x > 1$ we have in turn:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{1} f(t)dt + \int_{1}^{x} f(t)dt = F(1) + \int_{1}^{x} \frac{2}{t^3}dt = 0 + \left[-\frac{1}{t^2}\right]_1^x = 1 - \frac{1}{x^2}.$$

In order to find the expected value, similarly as in Example 2 we subdivide the domain of integration. Note that the second summand is an improper integral, which we compute by finding respective limit – the integrand is not a constant function equal everywhere to 0:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{1} xf(x)dx + \int_{1}^{\infty} xf(x)dx = \int_{-\infty}^{1} x \cdot 0dx + \int_{1}^{\infty} x \cdot \frac{2}{x^3}dx =$$

$$= 0 + \lim_{a\to\infty} \int_{1}^{a} \frac{2}{x^2}dx = \lim_{a\to\infty} [-2/x]_1^a = \lim_{a\to\infty} (-2/a + 2/1) = 2.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

In similar way we find the variance:

$$V(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - [E(X)]^2 = \int_{-\infty}^{1} x^2 f(x)dx + \int_{1}^{\infty} x^2 f(x)dx - [E(X)]^2 =$$
$$= \int_{-\infty}^{1} x^2 \cdot 0 dx + \int_{1}^{\infty} x^2 \cdot \frac{2}{x^3} dx - [E(X)]^2 = 0 + \lim_{a\to\infty} \int_{1}^{a} \frac{2}{x} dx - [E(X)]^2 =$$
$$= \lim_{a\to\infty} [2\ln|x|]_1^a = \lim_{a\to\infty} (2\ln|a| + 2\ln 1) = \infty.$$

It means that given variable has infinite variance (and standard deviation). In such a situation we say that the random variable does not have variance (or standard deviation).

Example 5
The probability that a customer of the financial intermediary will take out the loan is 0.3. The intermediary met 7 customers today. What is the probability that at least 2 of them will take out the loan?

Solution
For each customer the probability of taking out the loan is the same, equal to $p = 0.3$. In such a situation we can identify the number of customers taking out the loan with the binomially distributed random variable X. We make $n = 7$ trials (the total number of customers), and we are interested in the probability that $X \geq 2$. In order to find it we can add the probabilities that X takes a value between 2 and 7:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7).$$

We can also simplify the calculations. Note that the opposite of "$X \geq 2$" is "$X < 2$", thus the following holds:

$$P(X \geq 2) = 1 - P(X < 2).$$

The latter one is equal to:

$$P(X < 2) = P(X = 0) + P(X = 1).$$

Finally:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1).$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Recall that for the variable $X \sim B(n, p)$, the probability of occurring k successes is:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Hence finally:

$$\begin{aligned} P(X \geq 2) &= 1 - \binom{7}{0} \cdot 0.3^0 \cdot 0.7^7 - \binom{7}{1} \cdot 0.3^1 \cdot 0.7^6 \\ &= 1 - \frac{7!}{7!0!} \cdot 0.3^0 \cdot 0.7^7 - \frac{7!}{6!1!} \cdot 0.3^1 \cdot 0.7^6 = \\ &= 1 - 1 \cdot 0.3^0 \cdot 0.7^7 - 7 \cdot 0.3^1 \cdot 0.7^6 \approx 1 - 0.082 - 0.247 = 0.671. \end{aligned}$$

As we can see, the probability that at least two customers will take the loan, equals to 0.671.

Example 6
There are on average 3 customers by the counter in the bank. What is the probability that after entering this bank you will see at most one person in the queue? What is the probability that there will be at least 5 customers?

Solution
We know only the average number of customers (i.e., the average number of successes), so we assume that the number of customers is $X \sim Po(3)$. The probability that this number is not greater than 1 equals to:

$$P(X \leq 1) = P(X = 0) + P(X = 1).$$

We know that the probability of taking the value k by a variable having Poisson distribution is equal to:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

It follows that the searched probability equals to:

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{3^0}{0!} e^{-3} + \frac{3^1}{1!} e^{-3} \approx 0.050 + 0.149 = 0.199.$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

Similarly as in the previous example, the probability that $X \geq 5$ can be derived from the relation: $P(X \geq 5) = 1 - P(X < 5)$:

$$P(X \geq 5) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) =$$
$$= 1 - \frac{3^0}{0!}e^{-3} - \frac{3^1}{1!}e^{-3} - \frac{3^2}{2!}e^{-3} - \frac{3^3}{3!}e^{-3} - \frac{3^4}{4!}e^{-3} \approx$$
$$\approx 1 - 0.050 - 0.149 - 0.224 - 0.224 - 0.168 = 0.185.$$

Example 7
The salaries of the employees of the company have distribution $N(2{,}500, 1{,}000)$. The holiday funding for holiday from the social benefits fund is granted to people with incomes not exceeding 1,500 PLN. What percentage of the crew will receive the funding?

Solution
Let us denote by X the random variable equal to the salary of randomly chosen employee. It follows that $X \sim N(2{,}500, 1{,}000)$. We are looking for the probability that the this variable is not greater than 1,500: $P(X \leq 1{,}500)$. In order to find it, first we need to standardize the value. For $x = 1{,}500$, the standardized value is:

$$x_{st} = (x - \mu)/\sigma = (1{,}500 - 2{,}500)/1{,}000 = -1.00$$

(remember to round to the nearest 0.05).
In the table of the CDF of the standard normal distribution (Section 14.3) for $x = -1.00$ we find the probability 0.159. It means that approximately 15.9% of the employees will receive the funding.

Example 8
The probability that a customer of the financial intermediary will take out the loan is 0.3. The intermediary met 200 customers this month. They will receive a bonus if at least 50 customers will take out the loan. What is the probability this event?

Solution
This example is very similar to Example 5. Theoretically, it could be solved in the same way, but the problem is the number of trials $n = 200$. The determination of the factorial of such a number exceeds the capacity of many calculators. Furthermore, adding the 50 small numbers can cause the error approximations to cumulate and really affect the result. Therefore, we approximate the binomial distribution by another one. Since $p = 0.3 > 0.1$ (and moreover $np = 60 > 20$), we use the approximation by normal distribution. In

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

order to determine the probability $P(X \in [a, b])$, we have to find the probability $P(X \in [a - 0.5, b + 0.5])$, thus in this case we have to find the probability $P(X \in [50 - 0.5, 200 + 0.5]) = P(X \in [49.5, 200.5])$. We have to standardize the ends of the interval. Since $\mu = np = 60$ and $\sigma = \sqrt{np(1-p)} \approx 6.481$, we have $a_{st} \approx (49.5 - 60)/6.481 \approx -1.62$ and $b_{st} \approx (200.5 - 60)/6.481 \approx 21.68$. After rounding it to the nearest 0.05 we obtain –1.60 and 21.70. The probabilities found in the table of normal distribution are respectively 0.055 and 1 (for $x \geq 4$ we assume that $\Phi(x) = 1$), so the searched probability equals to: $P(X \in [49.5, 200.5]) \approx 1 - 0.055 = 0.945$. The intermediary may therefore be almost sure of receiving the bonus.

Example 9
The probability that a customer of the financial intermediary will take out the loan is 0.05. The intermediary met 200 customers this month. They will receive a bonus if at least 5 customers will take out the loan. What is the probability this event?

Solution
Also in this case we approximate the binomial distribution by another one. Since $p = 0.05 < 0.1$ and $np = 10 < 20$, we use the Poisson distribution.
We have $\lambda = np = 10$, hence the result is:

$$P(X \geq 5) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) =$$
$$= 1 - \frac{10^0}{0!}e^{-10} - \frac{10^1}{1!}e^{-10} - \frac{10^2}{2!}e^{-10} - \frac{10^3}{3!}e^{-10} - \frac{10^4}{4!}e^{-10} \approx$$
$$\approx 1 - 0.000 - 0.000 - 0.002 - 0.008 - 0.019 = 0.971.$$

Also in this case, the intermediary may be almost sure of receiving the bonus.

EXERCISES

1. Find the CDF, expectation and standard deviation of the number obtained in single rolling of a dice.
2. Find the CDF, expectation and standard deviation of the number of heads obtained in flipping two coins.
3. Find the CDF, expectation and standard deviation of the random variable described by the density function:

a) $$f(x) = \begin{cases} \frac{1}{18}x, & \text{if} \quad x \in [0,6], \\ 0, & \text{if} \quad x \notin [0,6]; \end{cases}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

b) $f(x)=\begin{cases}\frac{1}{8}x, & \text{if} \quad x\in[0,4],\\ 0, & \text{if} \quad x\notin[0,4].\end{cases}$

4. Find the CDF, expectation and standard deviation of the random variable described by the density function:

a) $f(x)=\begin{cases}\frac{3}{x^4}, & \text{if } x\geq 1,\\ 0, & \text{if } x<1;\end{cases}$

b) $f(x)=\begin{cases}\frac{4}{x^5}, & \text{if } x\geq 1,\\ 0, & \text{if } x<1.\end{cases}$

5. The probability that a company taking out the loan will fail, is 0.25. Each client advisor supports 15 randomly selected companies. We choose a random advisor.
   a) What is the probability that exactly 4 of their customers will fail?
   b) What is the probability that at most 3 of their customers will fail?
   c) What is the probability that at least 2 of their customers will fail?
6. The daily demand for certain good (in pieces) has distribution $N(1{,}200, 400)$. We assume that the sale is conducted for 250 days a year. Estimate the number of days when:
   a) sales will be not less than 1,500;
   b) sales will be not less than 1,300;
   c) sales will not exceed 1,000.
7. There are on average 4 cars at a gas station on a busy highway. What is the probability that in a chosen moment:
   a) there will be exactly 4 cars at this gas station;
   b) there will be at most 2 cars at this gas station;
   c) there will be at least 3 cars at this gas station?
8. The probability that a company taking out the loan will fail, is 0.35. Each client advisor supports 200 randomly selected companies. We choose a random advisor.
   a) What is the probability that exactly 60 of their customers will fail?
   b) What is the probability that at most 50 of their customers will fail?
   c) What is the probability that at least 85 of their customers will fail?

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

9. The probability that a company taking out the loan will fail, is 0.02. Each client advisor supports 40 randomly selected companies. We choose a random advisor.
    a) What is the probability that exactly 5 of their customers will fail?
    b) What is the probability that at most 3 of their customers will fail?
    c) What is the probability that at least 2 of their customers will fail?

SOLUTIONS

1. Distribution: $\{(1, 1/6), (2, 1/6), (3, 1/6), (4, 1/6), (5, 1/6), (6, 1/6)\}$; $E(X) = 3.5$; $V(X) = 35/12$; $S(X) \approx 1.708$; $F(x) = 0$ for $x \in (-\infty, 1]$, $1/6$ for $x \in (1, 2]$, $1/3$ for $x \in (2, 3]$, $1/2$ for $x \in (3, 4]$, $2/3$ for $x \in (4, 5]$, $5/6$ for $x \in (5, 6]$, 1 for $x \in (6, \infty)$.

2. Distribution: $\{(0, 1/4), (1, 1/2), (2, 1/4)\}$; $E(X) = 1$; $V(X) = 0.5$; $S(X) \approx 0.707$; $F(x) = 0$ for $x \in (-\infty, 0]$, 0.25 for $x \in (0, 1]$, 0.75 for $x \in (1, 2]$, 1 for $x \in (2, \infty)$.

3. a) $E(X) = 4$; $V(X) = 2$; $S(X) \approx 1.414$; $F(x) = 0$ for $x \in (-\infty, 0]$, $1/36x^2$ for $x \in (0, 6]$, 1 for $x \in (6, \infty)$;
   b) $E(X) = 8/3$; $V(X) = 8/9$; $S(X) \approx 0.943$; $F(x) = 0$ for $x \in (-\infty, 0]$, $1/16x^2$ for $x \in (0, 4]$, 1 for $x \in (4, \infty)$.

4. a) $E(X) = 1.5$; $V(X) = 0.75$; $S(X) \approx 0.866$; $F(x) = 0$ for $x \in (-\infty, 1]$, $1 - 1/(x^3)$ for $x \in (1, \infty)$;
   b) $E(X) = 4/3$; $V(X) = 2/9$; $S(X) \approx 0.471$; $F(x) = 0$ for $x \in (-\infty, 1]$, $1 - 1/(x^4)$ for $x \in (1, \infty)$.

5. Distribution $B(15, 0.25)$;
   a) 0.225; b) 0.461; c) 0.92.

6. We multiply the probabilities by 250 and round to full days;
   a) 57; b) 100; c) 77.

7. Distribution $Po(4)$;
   a) 0.195; b) 0.091; c) 0.762.

8. Distribution $B(200, 0.35)$ approximated by $N(70, 6.745)$;
   a) 0.02; b) 0.002; c) 0.016.

9. Distribution $B(40, 0.02)$ approximated by $Po(0.8)$;
   a) 0.001; b) 0.99; c) 0.192.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343

14.3. APPENDIX: CUMULATIVE DISTRIBUTION OF $N(0, 1)$

In order to read the value of CDF for chosen x, we begin with rounding x to the nearest 0.05. For example $x = 1.7734$ is rounded to 1.75. Now we are looking for the respective probability in Table 14.3. For this purpose we find the number to the left of the decimal point in the column headers, and the digits after the decimal point – in the row headers. In our example these are respectively the column denoted by "1." and the row denoted by ".75". Their common element is the number 0.960 and this is the searched probability. Similarly if we were looking for the probability for –2.33, we should round it to –2.35, and then read the number in the column "–2." and row ".35": 0.009.
Assume that for all $x \geq 4$ the value of the CDF equals to 1, while for $x \leq -4$ it equals to 0.

Table 14.3. Cumulative distribution function of $X \sim N(0, 1)$: $\Phi(x) = P(X < x)$

	–3.	–2.	–1.	–0.	0.	1.	2.	3.
.00	0.001	0.023	0.159	0.500	0.500	0.841	0.977	0.999
.05	0.001	0.020	0.147	0.480	0.520	0.853	0.980	0.999
.10	0.001	0.018	0.136	0.460	0.540	0.864	0.982	0.999
.15	0.001	0.016	0.125	0.440	0.560	0.875	0.984	0.999
.20	0.001	0.014	0.115	0.421	0.579	0.885	0.986	0.999
.25	0.001	0.012	0.106	0.401	0.599	0.894	0.988	0.999
.30	0.000	0.011	0.097	0.382	0.618	0.903	0.989	1.000
.35	0.000	0.009	0.089	0.363	0.637	0.911	0.991	1.000
.40	0.000	0.008	0.081	0.345	0.655	0.919	0.992	1.000
.45	0.000	0.007	0.074	0.326	0.674	0.926	0.993	1.000
.50	0.000	0.006	0.067	0.309	0.691	0.933	0.994	1.000
.55	0.000	0.005	0.061	0.291	0.709	0.939	0.995	1.000
.60	0.000	0.005	0.055	0.274	0.726	0.945	0.995	1.000
.65	0.000	0.004	0.049	0.258	0.742	0.951	0.996	1.000
.70	0.000	0.003	0.045	0.242	0.758	0.955	0.997	1.000
.75	0.000	0.003	0.040	0.227	0.773	0.960	0.997	1.000
.80	0.000	0.003	0.036	0.212	0.788	0.964	0.997	1.000
.85	0.000	0.002	0.032	0.198	0.802	0.968	0.998	1.000
.90	0.000	0.002	0.029	0.184	0.816	0.971	0.998	1.000
.95	0.000	0.002	0.026	0.171	0.829	0.974	0.998	1.000

Plik zabezpieczony watermarkiem jawnym i niejawnym: 11274643A6432343